



UNIVERSIDAD NACIONAL DE COLOMBIA

SEDE MANIZALES

Sistema de clasificación basado en técnicas inteligentes para identificar el grado de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia

Rodolfo Mosquera Navarro

Universidad Nacional de Colombia
Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Industrial
Manizales, Colombia
Febrero 26 de 2021

Sistema de clasificación basado en técnicas inteligentes para identificar el grado de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia

Rodolfo Mosquera Navarro

Tesis presentada como requisito parcial para optar al título de:
Doctor en Ingeniería - Industria y Organizaciones

Director (a):

Ph.D., Omar Danilo Castrillón Gómez

Codirector (a):

Ph.D., Liliana Parra Osorio

Línea de Investigación:

Organizaciones, sistemas y gestión de la tecnología, la Información, el conocimiento y la innovación tecnológica.

Universidad Nacional de Colombia

Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Industrial

Manizales, Colombia

Febrero 26 de 2021

No he fracasado. He encontrado 10000 soluciones que no funcionan. Thomas Edison

A Dios, a mi madre por su constante amor y apoyo, a mis directores, siempre por su constante apoyo, ayuda y dedicación.

Agradecimientos

A Dios, por ser mi camino, verdad y vida.

A mi Madre Magnolia, por ser mi ejemplo e inspiración y lucha, para la consecución de todos mis logros.

A mis asesores, Omar y Liliana, que de una u otra forma me apoyaron con sus conocimientos y tiempo, para que mi proyecto se hiciera realidad.

Al profesor Andrés Carrión por su ayuda, amistad y disposición de su grupo de investigación en mi estancia doctoral en la Universidad Politécnica de Valencia, España.

A todos aquellos docentes y personas quienes estuvieron dispuestos a participar en el proceso de toma de datos, disponiendo de su tiempo y confort.

A MinCiencias (anterior Colciencias) por su programa de créditos condonables, convocatoria 647 para adelantar estudios doctorales a nivel nacional.

Este trabajo se realiza en el marco de:

“Convocatoria Nacional para el Apoyo al Desarrollo de Tesis de Posgrado o de Trabajos Finales de Especialidades en el área de la Salud de la Universidad Nacional de Colombia 2017-2018” Resolución 21 de 2017. Oficina de la Vicerrectoría de Investigación (21 de diciembre de 2017), proyecto seleccionado, financiado y ejecutado con recursos de la Universidad Nacional de Colombia con el número de identificación 40976 en el Sistema de Información Hermes. Anexo A.

Resumen

Sistema de clasificación basado en técnicas inteligentes para identificar el grado de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia

El objetivo del presente trabajo es el desarrollo de un algoritmo inteligente para mejorar la predicción del riesgo psicosocial entre los docentes de colegios públicos en Colombia. El enfoque está compuesto por el modelo de redes neuronales artificiales vinculado a la teoría física de la tensión superficial en los líquidos. Para lograr los objetivos de este estudio, los docentes de colegios públicos han completado la evaluación de la Batería para la evaluación de factores de riesgo psicosocial intralaboral para la identificación del nivel de riesgo. Las variables que componen los factores de riesgo psicosocial se utilizan como entradas y el nivel de riesgo se utiliza como salida en el algoritmo. La eficiencia de la red neuronal de tensión superficial física (RNA-TS) se examina contra los algoritmos árboles de decisión (algoritmo J48), Naïve Bayes (NBC), red neuronal artificial (ANN), máquina de vectores de soporte (SVM), máquinas de vectores de soporte con función de base radial (SVM-RBF), máquina de vector de soporte con escalada de colinas (HC-SVM), el algoritmo híbrido de búsqueda de cuco junto a máquinas vectoriales de soporte (CS-SVM) y el modelo híbrido de máquina de vector de soporte con k-vecino más cercano (k-NN-SVM), utilizando el nivel de sensibilidad, la especificidad, el nivel de exactitud, el porcentaje de error de clasificación y la curva ROC. Los resultados concluyeron que RNA-TS proporciona una mejor clasificación que los demás algoritmos. Por tanto, la RNA-TS se utiliza para predecir y clasificar el nivel de riesgo psicosocial de los docentes de colegios públicos en su actividad laboral. Uno de los temas importantes para aplicar la teoría de la tensión superficial en el entorno del mundo real es desarrollar un modelo que soporte el modelo de aprendizaje automático para reflejar toda la complejidad de los factores psicosociales del mundo en los entornos laborales y permitir su predicción. Los profesores que presentan un riesgo psicosocial muy alto y alto son identificados con un 97,37% de exactitud. Esto

ayudará a los gerentes a prever si los trabajadores están satisfechos con su carga de trabajo en el contexto de la higiene y la seguridad laboral. Finalmente, este es el primer estudio que introduce un algoritmo de tensión superficial física adaptado como clasificador inteligente para la predicción eficiente de factores de riesgo psicosocial en docentes de colegios públicos en sistemas académicos.

Palabras clave: Red Neuronal artificial de retropropagación, tensión superficial física, riesgo psicosocial, docentes de colegios públicos, Colombia.

Abstract

Classification system to the predicting of psychosocial risk level on state-school teachers in Colombia based on Artificial Intelligence

The main goal of this research is to develop an intelligent algorithm to improve the prediction of psychosocial risk among the state-schools teachers in Colombia. The model is composed for artificial neural network backpropagation linked to the physical theory of surface tension in liquids. To achieve the goals of this study, state school teachers have carried out the evaluation of the battery for the evaluation of intra-occupational psychosocial risk factors to identify the risk level. The variables that make up the psychosocial risk factors are used as inputs and the risk level is used as output in the algorithm. The efficiency of the physical surface tension neural network (PST-NN) is examined against decision trees algorithm (algorithm J48), Naïve Bayes (NBC), artificial neural network (ANN), support vector machine (SVM), support vector machines with radial basis function (SVM-RBF), hill climbing random restart - support vector machine hybrid model (HCRR-SVM), cuckoo search - support vector machines hybrid algorithm (CS-SVM) and the k-nearest neighbor – support vector machine hybrid model (k-NN-SVM), for metric evaluation were used, the sensitivity level, specificity, accuracy level, percentage of classification error, and ROC curve. The results concluded that PST-NN provides a better classification than the other algorithms. Therefore, the PST-NN is used to predict and classify the level of psychosocial risk of state-school teachers in their work activity. One of the important topics for applying the theory of surface tension in the real world environment, is to develop a model that supports the machine learning model to reflect all the complexity of the psychosocial factors of the world in work environments and allow its prediction. Teachers who present a very high and high psychosocial risk level were identified with 97.37% accuracy. This will help managers to predict whether workers are satisfied with their workload in the context of occupational hygiene and safety. Finally, this is the first study that introduces a physical surface tension algorithm adapted as an intelligent classifier for the efficient prediction of psychosocial risk factors in state-school teachers in academic systems.

Keywords: Artificial Neural Network backpropagation, physical surface tension, Classification, psychosocial risk, state-school teachers, Colombia.

Contenido

PÁG.

I. ARTÍCULOS PUBLICADOS EN REVISTA DE ALTO IMPACTO Y ARTÍCULOS SOCIALIZADOS EN CONGRESOS	21
ARTÍCULOS EN REVISTA INDEXADA Q1 Y A1.....	21
ARTÍCULOS EN REVISTA INDEXADA A2.....	22
ARTÍCULOS EN CONGRESOS INTERNACIONALES	22
PONENCIAS EN CONGRESOS INTERNACIONALES.....	22
CONGRESOS NACIONALES.....	23
PUBLICACIONES QUE HAN REALIZADO CITAS A LOS ARTÍCULOS PUBLICADOS SOBRE LA TESIS DE INVESTIGACIÓN.....	24
II. CAPÍTULO 1 – PLANTEAMIENTO GENERAL	31
1.1. VACÍO DEL CONOCIMIENTO	31
1.1.1. <i>Situación problemática a abordar en la investigación</i>	37
1.2.1. <i>Relación factores de riesgo psicosocial e inteligencia artificial</i>	41
1.2. SISTEMATIZACIÓN DEL PROBLEMA.....	45
1.3. JUSTIFICACIÓN.....	46
1.4. DISEÑO GENERAL DE LA INVESTIGACIÓN	50
1.5. HIPÓTESIS GENERAL DE LA INVESTIGACIÓN.....	51
1.5.1. <i>Identificación y definición conceptual y operacional de las variables</i>	51
1.5.2. <i>Comprobación de la hipótesis de investigación</i>	52
1.5.3. <i>Tipo de investigación</i>	53
1.6. OBJETIVO GENERAL.....	53
1.7. OBJETIVOS ESPECÍFICOS	53
III. CAPÍTULO 2 - MARCO TEÓRICO Y REFERENCIAL DE LA INVESTIGACIÓN.....	55
2.1. INTELIGENCIA ARTIFICIAL.....	56
2.2. APRENDIZAJE SUPERVISADO	58
2.	58
2.2.1. <i>Clasificación</i>	58
2.2.2 <i>Métodos de clasificadores lineales</i>	60
2.2.3 <i>Redes Neuronales Artificiales</i>	60
2.2.4 <i>Métricas de evaluación del desempeño</i>	64
2.2.5 <i>Curva ROC</i>	65
2.3. APLICACIÓN DE LAS TÉCNICAS DE INTELIGENCIA ARTIFICIAL EN EL CAMPO DE LA SALUD Y DE LA SEGURIDAD Y SALUD EN EL TRABAJO	66
2.4. TEORÍA DE LA TENSIÓN SUPERFICIAL	76

2.5.	FACTORES DE RIESGO PSICOSOCIAL	78
2.5.1.	<i>Modelo de demanda, control, apoyo social</i>	78
2.5.2.	<i>Modelo de desequilibrio esfuerzo-recompensa</i>	79
2.5.3.	<i>Modelo dinámico de los factores de riesgo psicosocial</i>	80
2.5.4.	<i>Condiciones Intralaborales de los riesgos psicosociales</i>	80
2.5.5.	<i>Conjunto de datos utilizado para este trabajo</i>	82
IV.	CAPÍTULO 3 - MATERIALES Y MÉTODOS	89
3.1	FACTORES DE RIESGO PSICOSOCIAL Y PROCESO DE CONSTRUCCIÓN DEL MODELO .	89
3.2.	BASE DE DATOS.....	91
3.2.1	<i>Factores de riesgo psicosocial, variables fisiológicas y musculoesqueléticas</i>	91
3.3.	PRE PROCESAMIENTO DE LOS DATOS.....	91
3.3.1.	<i>Conjunto de datos</i>	91
3.3.2.	<i>Análisis de datos</i>	95
3.4.	DESARROLLO DE ALGORITMOS PREVIOS A LA OBTENCIÓN DEL MODELO PROPUESTO	98
3.4.1.	<i>Algoritmo de ascenso de colinas con reinicio aleatorio – Máquinas de Soporte Vectorial (RRHC-SVM)</i>	99
3.4.2.	<i>Algoritmo de búsqueda cuckoo – Máquinas de Soporte Vectorial (CS-SVM)</i> 100	
3.5.	PASO 1. EL ALGORITMO PROPUESTO DE RED NEURONAL DE TENSIÓN SUPERFICIAL (RNA-TS)	102
3.5.1.	<i>Conversión de variables matemáticas a variables físicas</i>	105
3.5.2.	<i>Inicialización de parámetros de red neuronal</i>	106
3.6.	PASO 2. ENTRENAMIENTO DE LA RED NEURONAL ARTIFICIAL (PST-NN).....	107
3.6.1.	<i>Cálculo de salida de la capa 1</i>	107
3.6.2.	<i>Cálculo de salida de la capa 2</i>	108
3.6.3.	<i>Cálculo de salida de la capa 3</i>	109
3.6.4.	<i>Cálculo de salida de la capa de tensión superficial</i>	110
3.6.5.	<i>Error de Retropropagación</i>	113
3.6.6.	<i>La derivada del error, con respecto a la salida de la red neuronal</i>	114
3.6.7.	<i>La derivada del error, con respecto a la salida de la capa 3</i>	114
3.6.8.	<i>Derivada del error, según los pesos de la capa 3</i>	115
3.6.9.	<i>Derivada del error, de acuerdo con el sesgo de la capa 3</i>	115
3.6.10.	<i>Derivada del error, con respecto a la salida de la capa 2</i>	115
3.6.11.	<i>Derivada del error, según los pesos de la capa 2</i>	116
3.6.12.	<i>Derivada del error, de acuerdo con el sesgo de la capa 2</i>	116
3.6.13.	<i>Derivada del error, con respecto a la salida de la capa 1</i>	116
3.6.14.	<i>Derivada del error, según los pesos de la capa 1</i>	117
3.6.15.	<i>Derivada del error, de acuerdo con el sesgo de la capa 1</i>	117
3.6.16.	<i>Criterios para la parada de entrenamiento</i>	118
3.6.17.	<i>Medidas estadísticas para la prueba de clasificación</i>	119
3.7.	PASO 3. ANÁLISIS DE ESTABILIDAD.....	120
3.8.	PASO 4. CLASIFICACIÓN DEL NIVEL DE RIESGO PSICOSOCIAL Y RECOMENDACIONES PARA SU PREVENCIÓN	121
3.9.	COMPARACIÓN CON OTRAS TÉCNICAS INTELIGENTES	121

3.9.1.	<i>Comparación con el algoritmo j48 árbol de decisión y naive bayes</i>	121
3.9.2.	<i>Comparación con la red neuronal artificial de retro-propagación Perceptrón utilizando algoritmo genético para reducir la dimensionalidad de los datos</i>	122
3.9.3.	<i>Comparación con máquinas vectoriales (svm lineal)</i>	124
3.9.4.	<i>Comparación con máquinas de soporte vectorial Rbf y algoritmo híbrido ascenso de colina-máquinas de soporte vectorial (hc-svm)</i>	125
3.9.5.	<i>Comparación con el algoritmo híbrido knn-máquina de vectores de soporte (knn-svm)</i>	125
V.	CAPÍTULO 4 -RESULTADOS Y DISCUSIÓN – COMPROBACIÓN DE LA HIPÓTESIS DE LA INVESTIGACIÓN	126
4.2.	DESARROLLO DE ALGORITMOS PREVIOS A LA OBTENCIÓN DEL MODELO PROPUESTO	126
4.2.1.	<i>Algoritmo de ascenso de colinas con reinicio aleatorio – Máquinas de Soporte Vectorial (RRHC-SVM)</i>	126
4.2.2.	<i>Algoritmo de búsqueda cuckoo – Máquinas de Soporte Vectorial (CS-SVM)</i>	127
4.3.	PASO 1 Y 2. ALGORITMO DE RED NEURONAL DE TENSIÓN SUPERFICIAL FÍSICA.....	127
4.3.1.	<i>Masa(m) y Perímetro(Per)</i>	128
4.3.2.	<i>Rendimiento del modelo de red neuronal de tensión física superficial</i>	129
4.3.3.	<i>Matriz de confusión para el modelo de Red neuronal de tensión superficial</i>	131
4.3.4.	<i>Curva ROC</i>	132
4.4.	PASO 3. ANÁLISIS DE ESTABILIDAD	134
4.5.	PASO 4. CLASIFICACIÓN DEL NIVEL DE RIESGO PSICOSOCIAL Y RECOMENDACIONES PARA SU PREVENCIÓN.....	135
4.6.	RESULTADOS EXPERIMENTALES BASADOS EN MÉTODOS DE COMPARACIONES PREVIAS.....	136
4.7.	DISCUSIÓN DE RESULTADOS	137
4.7.1.	<i>Aspectos generales</i>	137
4.7.2.	<i>Aspectos específicos</i>	140
VI.	5. CONCLUSIONES, FUTURAS LÍNEAS DE INVESTIGACIÓN Y LIMITACIONES	149
5.1.	CONCLUSIONES	149
5.2.	FUTURAS LÍNEAS DE INVESTIGACIÓN	153
5.3.	LIMITACIONES	154
VII.	REFERENCIAS	156
VIII.	6. ANEXO A: PROYECTO DE INVESTIGACIÓN EJECUTADO	178
IX.	ANEXO B: PASANTÍA INVESTIGATIVA	179
X.	ANEXO C: ESTRATEGIA DE BÚSQUEDA SISTEMÁTICA, ECUACIONES DE BÚSQUEDA, RESULTADOS OBTENIDOS Y DISTRIBUCIÓN DE PUBLICACIONES POR AÑO	180
XI.	ANEXO D: PROTOCOLO PARA LA APLICACIÓN DE LA BATERÍA DE RIESGO PSICOSOCIAL Y PARA LA TOMA DE SEÑALES Y FORMATO DE CONSENTIMIENTO INFORMADO	183

XII. ANEXO E: INTERFAZ PARA LA GESTIÓN DEL RIESGO PSICOSOCIAL BASADO EN EL ALGORITMO DE RED NEURONAL DE TENSION SUPERFICIAL	186
XIII. ANEXO F: CÓDIGO ALGORITMO DE ASCENSO DE COLINAS CON REINICIO ALEATORIO – MÁQUINAS DE SOPORTE VECTORIAL (RRHC-SVM).....	187
XIV. ANEXO G: CÓDIGO ALGORITMO DE BÚSQUEDA CUCKOO – MÁQUINAS DE SOPORTE VECTORIAL (CS-SVM).....	192
XV. ANEXO H: CÓDIGO ALGORITMO RED NEURONAL DE TENSION SUPERFICIAL (RNA-TS).....	196
XVI. ANEXO I: CARTA DE CERTIFICACIÓN POR PARTE DEL PSICÓLOGO LABORAL QUE COLABORÓ EN EL PROYECTO EN LA TOMA DE DATOS.....	203
XVII. ANEXO J: PROYECTO DE INVESTIGACIÓN LLEVADO ACTUALMENTE POR EL INVESTIGADOR EN EL TEMA DE LA TESIS DOCTORAL.....	204

Lista de figuras

PÁG.

Figura 0 - 1. Esquema de la introducción.....	26
Figura 1- 1. Identificación de síntomas de educación básica primaria y secundaria en colegios públicos de Colombia. Adaptado de: Espinoza (2015); García (2010); Diazgranados (2006), Restrepo (2011).....	43
Figura 1- 2. Vacío del conocimiento.....	45
Figura 1- 3. Hilo conductor de la investigación.....	50
Figura 1- 4. Operacionalización de variables.	52
Figura 2 - 1. Esquema del marco teórico y referencial de la investigación.	55
Figura 2- 2. Jerarquía de diferentes algoritmos de aprendizaje automático que incluyen técnicas de aprendizaje supervisadas versus no supervisadas versus reforzadas.	59
Figura 2- 3. Fuerzas entre dos moléculas de un líquido.....	77
Figura 2- 4. Modelo teórico de exigencias psicológicas- demandas-control apoyo social	79
Figura 3 - 1. Pasos para la construcción del modelo para la identificación del nivel de riesgo psicosocial en docentes de colegios públicos.	90
Figura 3 - 2. Aplicación de la batería de riesgos intralaboral a cada docente y evaluación de condiciones fisiológicas a docentes con problemas médicos declarados.	95
Figura 3 - 3. Aplicación algoritmo ReliefF para reducción de la dimensionalidad de los datos.	97
Figura 3 - 4. Flujograma de la metodología propuesta.....	101
Figura 3 - 5. Strider de agua sobre el líquido debido a la tensión superficial.....	103
Figura 3 - 6. Flujograma del algoritmo RNA -TS propuesto.....	103
Figura 3 - 7. Red Neuronal de tensión superficial física (RNA-TS).....	106
Figura 3 - 8. Red neuronal de tensión superficial.....	113
Figura 4 - 1. Modelo de clasificación del riesgo perímetro vs. Masa.	128
Figura 4 - 2. Visualización de la separación del nivel de riesgo en el modelo.	129
Figura 4 - 3. Rendimiento de iteraciones en el modelo de red neuronal de tensión superficial.....	130
Figura 4 - 4. Curva ROC para la clasificación del riesgo psicosocial del modelo Red neuronal de tensión superficial.	132
Figura 4 - 5. Flujo de recomendación para prevención del riesgo psicosocial basado en el algoritmo Red neuronal de tensión superficial.	135

Lista de tablas

PÁG.

Tabla 2 - 1. Matriz de confusión.....	64
Tabla 2 - 2. Variables de los dominios de factores de riesgo psicosocial intralaboral.....	83
Tabla 2 - 3. Variables musculo esqueléticas evaluadas. Tomado de: (Sauter & Murphy, 1984).	87
Tabla 2 - 4. Variables fisiológicas evaluadas. Tomado de: (Sauter & Murphy, 1984).....	88
Tabla 3 - 1. Variables para factores de riesgo intralaboral. Tomado de: (Villalobos et al., 2010).	92
Tabla 3 - 2. Variables fisiológicas. Tomado de: (Sauter & Murphy, 1984).	93
Tabla 3 - 3. Síntomas musculoesqueléticos.....	94
Tabla 3 - 4. Tensión superficial de líquidos (Jasper, 1972).	111
Tabla 3 - 5. División de datos de entrenamiento, validación y testeo.	119
Tabla 4 - 1. Resultados para el conjunto de datos.	130
Tabla 4 - 2. Matriz de confusión del modelo de Red neuronal de tensión superficial para la predicción del riesgo psicosocial, para el conjunto de datos de entrenamiento (72%), validación (8%) y prueba (20%).	131
Tabla 4 - 3. Medidas estadísticas para la prueba de clasificación para los cuatro niveles de riesgo. Para el conjunto de datos de entrenamiento (72%), validación (8%) y prueba (20%).	133
Tabla 4 - 4. Porcentaje de exactitud para resultados predictivos sobre el nivel de riesgo psicosocial en maestros de escuela colombianos.	134
Tabla 4 - 5. Análisis de varianza (Anova).....	134
Tabla 4 - 6. Resultados aplicando diferentes técnicas de clasificación. *Exactitud.....	136

I. Artículos publicados en revista de alto impacto y artículos socializados en congresos

Algunas partes de esta tesis han sido divulgadas de forma previa en revistas especializadas indexadas de alto impacto y en congresos relacionadas con el área a nivel nacional e internacional.

Artículos en revista indexada Q1 y A1

Artículo aprobado.

Mosquera, Rodolfo., Castrillón, O.D., Parra Osorio, L., Novais, P., Oliveira, T., & Valencia, J. F. (2021). *Improving classification based on physical surface tension-neural net for the prediction of psychosocial-risk level in public school teachers*. PeerJ Computer Science. (In Press).

Mosquera, Rodolfo., Parra-Osorio, L., & Castrillón, O. D. (2016). *Metodología para la Predicción del Grado de Riesgo Psicosocial en Docentes de Colegios Colombianos utilizando Técnicas de Minería de Datos*. Información tecnológica, 27(6), 259–272. <https://doi.org/10.4067/S0718-07642016000600026>

Artículos en revista indexada A2

Mosquera, Rodolfo, Castrillón, O. D., & Parra, L. (2018a). *Support Vector Machines, Naïve Bayes Classifier and Genetic Algorithms for the Prediction of Psychosocial Risks in Teachers of Colombian Public Schools*. *Información tecnológica*, 29(6), 153–162. <https://doi.org/10.4067/S0718-07642018000600153>

Mosquera, Rodolfo, Castrillón, O. D., & Parra, L. (2018b). *Predicción de Riesgos Psicosociales en Docentes de Colegios Públicos Colombianos utilizando Técnicas de Inteligencia Artificial*. *Información tecnológica*, 29(4), 267–280. <https://doi.org/10.4067/S0718-07642018000400267>

Artículos en congresos internacionales

Mosquera, Rodolfo, Castrillón, O. D., & Parra-Osorio, L. (2019). *Aplicación del modelo híbrido k-nearest neighbors- Support Vector Machine para la predicción del riesgo psicosocial en docentes de colegios públicos colombianos*. *Proceedings of the 17th Latin American and Caribbean Conference for Engineering and Technology*, 1, 5.

Mosquera, Rodolfo, Gómez, O. D. C., Osorio, L. P., & García, A. C. (2018). *Classification system for the predicting of psychosocial risk level in public-school teachers based on Artificial Intelligence*. *Proceedings of XVIII Conferencia de La Asociación Española Para la Inteligencia Artificial*, 1, 1367–1372.

Ponencias en Congresos internacionales

- Mosquera, Rodolfo, Castrillón, O. D., & Parra-Osorio, L. (2019). *Aplicación del modelo híbrido k-nearest neighbors- Support Vector Machine para la predicción del riesgo psicosocial en docentes de colegios públicos colombianos*. *Proceedings of the 17th Latin American and Caribbean Conference for Engineering and Technology*, 1, 5.

- Mosquera, Rodolfo, Gómez, O. D. C., Osorio, L. P., & García, A. C. (2018). *Classification system for the predicting of psychosocial risk level in public-school teachers based on Artificial Intelligence*. Proceedings of XVIII Conferencia de La Asociación Española Para La Inteligencia Artificial, 1, 1367–1372.
- Mosquera, R; Castrillon, G & Ruiz, S. (2015). *Aplicación de una metodología de toma de decisiones basada en Sistemas Expertos para la identificación de los peligros y la valoración y control de los riesgos en seguridad y salud ocupacional*, XV International Conference on Occupational Risk Prevention 2015, ID: 26340. Santiago de Chile, 23 al 26 de octubre de 2015.

Congresos nacionales

- Mosquera, R. (2017). *Identificación del grado de riesgo psicosocial en docentes de colegios públicos utilizando un clasificador naive bayes y algoritmos genéticos*. IV Coloquio de Investigación Posgrado Ingeniería Industrial. Manizales, 25 y 26 de octubre de 2017.
- Mosquera, R, Castrillon, OD & Parra, L. (2016). *Análisis y Evaluación del grado de riesgo psicosocial en docentes de colegios colombianos utilizando árboles de decisión*. XXXVI Congreso de Ergonomía, Higiene, Medicina y Seguridad Ocupacional, 15º Congreso Colombiano de Ergonomía: La Ergonomía del Mundo en Colombia. Hotel Intercontinental de Medellín, Medellín, noviembre 1, 2, 3 y 4 de 2016.
- Mosquera, R. (2016). *Análisis y Evaluación del grado de riesgo psicosocial en docentes de colegios colombianos utilizando técnicas de minería de datos*. III Coloquio de Investigación Posgrados de Ingeniería Industrial. Manizales, octubre 26, 27 y 28 de 2016.

- Mosquera, R. (2015). *Identificación de Riesgo Psicosocial desde un enfoque de ergonomía en docentes universitarios utilizando técnicas de Inteligencia Artificial*. II Coloquio de Investigación Posgrados de Ingeniería Industrial, Manizales, 3 y 4 de noviembre de 2015.

Publicaciones que han realizado citas a los artículos publicados sobre la tesis de investigación.

Varios autores han utilizado la información recabada con respecto a el conjunto de datos recopilado para hacer desarrollos propios y análisis de la información y han citado también los diferentes artículos que se han ido publicando a lo largo del desarrollo de ésta tesis doctoral, entre ellos tenemos:

- Vásquez León, C. A. (2020). Solución de minería de datos en la fidelización de agremiados del centro de capacitación del Colegio de Ingenieros del Perú sede La Libertad.
- Rivera, E., Moreno, R., Pérez, H., & Nakano, M. (2020). Separación de señales usando análisis de componentes principales y muestreo compresivo con mediciones mínimas. *Información tecnológica*, 31(1), 287-300.
- Galindo, E. A., Perdomo, J. A., & Figueroa-García, J. C. (2020). Estudio comparativo entre máquinas de soporte vectorial multiclase, redes neuronales y sistema de inferencia neurodifuso auto organizado para problemas de clasificación. *Información Tecnológica*, 31(1).
- Álvarez Obregón, M. A., Buitrago Cáceres, D. A., Roa Ruiz, M. F., & Tapia Ceballos, J. J. (2020). Diseño de un aplicativo para la identificación del estado de riesgo biopsicosocial en madres gestantes.

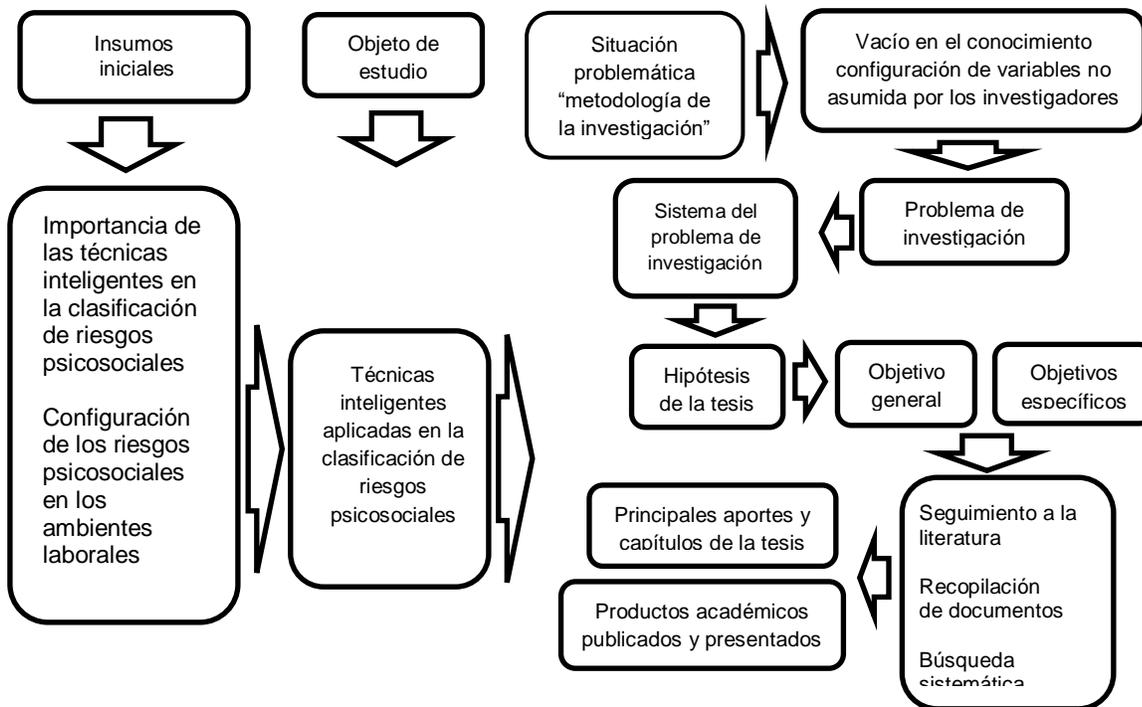
- Arrigoni, f. s. (2020). Burnout en personal docente y no docente de una escuela albergue de Mendoza (Argentina). *Perspectivas en Psicología*, 17(1), 32-42.
- Zuniga-Jara, S., & Pizarro-Leon, V. (2018). Mediciones de estrés laboral en docentes de un colegio público regional chileno. *Información tecnológica*, 29(1), 171-180.
- Agudelo-Torres, J. F., Rojas-Restrepo, F. S., Ocampo-Ruiz, E., & Clavijo-Zapata, S. J. (2018). Sobre la Evaluación Escolar y su Ética. *Información tecnológica*, 29(5), 71-80.
- Mina, M. A. E. (2018). Weka, áreas de aplicación y sus algoritmos: una revisión sistemática de literatura. *Revista Científica ECOCIENCIA*, 5, 1-26.
- Leon Atiquipa, H. E. (2018). Desarrollo de un modelo algorítmico basado en árboles de decisión para la predicción de la permanencia de un paciente en un proceso psicoterapéutico.
- Santamaría, C. N. V., Angulo, A. S., & Mariño, P. A. (2017). Análisis del diagnóstico de enfermedades mentales en la ciudad de Bogotá, mediante técnicas de minería de datos. *Mundo FESC*, (13), 35-47.
- Garcés, D. A., & Castrillón, O. D. (2017). Diseño de una Técnica Inteligente para Identificar y Reducir los Tiempos Muertos en un Sistema de Producción. *Información tecnológica*, 28(3), 157-170.

Introducción

El presente trabajo de investigación pretende desarrollar un sistema de clasificación basado en técnicas inteligentes que permitan y resulten útiles respecto a la predicción del grado de riesgo psicosocial al cual se encuentran expuestos educadores de colegios públicos colombianos.

El problema de investigación y la hipótesis se derivaron a partir del hilo conductor que se expone en la **Figura 0 - 1**.

Figura 0 - 1. Esquema de la introducción.



Fuente: Elaboración propia

El proceso se desarrolló como se explica a continuación:

Debido al paro nacional de docentes de colegios públicos que se presentó en mayo 7 de 2015 y donde se realizaron demandas de los docentes del magisterio, entre ellas las relacionadas con las condiciones de seguridad laboral de los docentes como los riesgos ergonómicos y psicosociales (Fecode, 2015), y que derivado de ello el gobierno nacional a través del Ministerio de Educación reglamentó el Decreto 1655 de agosto 20 de 2015, en el cual se establecen las actividades de promoción y prevención dentro del Sistema de Gestión de la Seguridad y la Salud en el Trabajo para los educadores afiliados al magisterio nacional (Ministerio de Educación, 2015) en su artículo 2.4.4.3.3.4, hace un llamado a que en el país se puedan: “Diseñar y ejecutar programas para la prevención y control de enfermedades generadas por riesgos psicosociales”, “Diseñar programas que se orienten a la prevención y control de las patologías generadas por el estilo de vida de los educadores activos y por su entorno laboral”, a raíz de esto, la presente tesis doctoral se desarrolló en el área metropolitana de la ciudad de Cali (Colombia), y se decidió investigar sobre el sector educativo público (colegios públicos) de la región.

A partir de esto y de la importancia del sector educativo público se determinó el objeto de estudio como el “Sistema de clasificación de riesgos psicosociales en docentes de colegios públicos”. Se incluyeron en la investigación dos factores relevantes: inteligencia artificial y riesgos psicosociales y se incorporó la teoría de la tensión superficial como modelo a emular para el desarrollo de la nueva metodología.

Muchos de los sistemas inteligentes son esencialmente basados en técnicas heurísticas y principios de la investigación de operaciones. Estas técnicas son muy variadas y pueden incluir, por ejemplo, la lógica difusa, redes neuronales artificiales, redes bayesianas, algoritmos evolutivos, que resultan eficientes en el modelado de sistemas complejos (Santos, 2011) y actualmente las técnicas inteligentes tienen muchos y muy diversos usos, como por ejemplo, en finanzas, en la educación, en el control del tráfico, en los sistemas de salud, en la robótica, en el control vehicular, entre otros (Chakraborty, 2020).

Las técnicas de inteligencia computacional o computación suave facilitan el desarrollo de proyectos de investigación o de aplicaciones tecnológicas novedosas en la toma de

decisiones y algunas técnicas como la de explotación de la información permiten identificar las condiciones para encontrar un resultado (atributo clase) ya definido el dominio del problema, donde se puede reconocer la relación entre los datos y definir la relación entre atributos (Salal et al., 2019).

Existe gran diversidad de proyectos desde la investigación en Sistemas Inteligentes que utilizan las técnicas para facilitar el análisis estadístico de la información y apoyar líneas de investigación en diversos frentes, como, por ejemplo, el diseño de evaluación por competencias de estudiantes de asignaturas en universidades, con el fin de mejorar el proceso de aprendizaje o predecir las variables que pueden afectar el desempeño de los estudiantes (Aldowah et al., 2019; Berens et al., 2019).

La tendencia general en la promoción de la Seguridad y la Salud en el Trabajo abarca no solo a los riesgos físicos, químicos, biológicos, eléctricos, mecánicos y biomecánicos, sino que incluye a los riesgos psicosociales, que influyen en los ambientes de trabajo empresarial y en el bienestar físico y mental del trabajador (Burke & Pignata, 2020).

La incidencia de los factores psicosociales en el trabajo contribuye al deterioro de la salud. Existen estudios que demuestran la relación entre síndromes no muy bien caracterizados de origen psicológico, expresados de forma somática y condiciones de trabajo estresantes. Los factores psicosociales han ido en aumento en todos los sectores laborales desfavorablemente y se transforman en una amenaza para el trabajador (Sauter & Murphy, 1984).

La creciente necesidad de desarrollar instrumentos o herramientas para la identificación de factores de riesgo psicosocial nace de la importancia y magnitud que tienen los riesgos psicosociales a nivel mundial en todas las organizaciones y las implicaciones en la salud de los trabajadores, en la productividad y específicamente en el bienestar laboral (Hernández et al., 2020; Kortum & Leka, 2014; Maqsoom et al., 2020).

Se ha identificado que el sector educativo es uno de los grupos de trabajo que desarrolla sintomatologías físicas y mentales que inciden en la salud debido al tipo de trabajo que desarrolla (Botero Alvarez, 2013), ya que éste tipo de actividad afecta al docente pues involucra situaciones personales propias de otras personas, lo que contribuye a desarrollar distrés (Ansoleaga et al., 2014; Botero Alvarez, 2013) y los resultados epidemiológicos muestran una relación entre el bienestar mental y los factores psicosociales presentes en el trabajo que de manera directa afectan el desempeño de los trabajadores (Villalobos F, 2004).

De manera que el trabajo académico implica riesgos específicos que perjudican tanto al docente como a las instituciones de educación. Se hace necesario llevar a cabo investigaciones sistemáticas que promocionen y promuevan la salud y seguridad en el trabajo, aspectos a los que las instituciones de educación no son ajenas y requieren prestar mayor interés para diseñar e implementar programas de atención, prevención e intervención de los problemas que afectan el quehacer docente y la calidad de los servicios (Silva Gutiérrez & Vicente Flores, 2015).

El sector de la educación en Colombia es uno de los sectores de la economía más expuestos a enfermedades de factor psicosocial (Ministerio de la Protección Social, 2006), debido a los procesos organizacionales propios de la labor docente. Según remarca (Posada-Quintero et al., 2020) en 2002, el Ministerio de Educación de Colombia emitió el estatuto 1278, de profesionalización docente, derogando el estatuto 2277 de 1977. Si bien el estatuto 1278 tenía como objetivo aumentar la calidad del servicio educativo y la remuneración de los docentes, existe evidencia de que las evaluaciones abundantes obstaculizaron la promoción de los docentes en sus cargos. El sistema introducido por el estatuto 1278 resultó en un deterioro de la calidad de vida de los maestros y una mayor incidencia de riesgo psicosocial.

El mismo (Posada Quintero et al., 2019) en un estudio anterior, había evaluado exactamente qué factores de riesgo y síntomas de agotamiento tenían mayor prevalencia para los maestros bajo el estatuto 1278, en comparación con los maestros bajo el estatuto

2277. Sin embargo, el estudio se limitó a las diferencias entre los grupos y no pudo extraer conclusiones sobre la importancia de los factores de riesgo y los síntomas. Los limitados recursos de las instituciones públicas y del sistema educativo en su conjunto exigen el análisis de la importancia de los factores de riesgo y síntomas psicosociales entre docentes de los dos estatutos. Identificar la importancia puede permitir que las autoridades institucionales y los responsables de la formulación de políticas asignen recursos para investigación a los temas más relevantes.

Por todo ello, el siguiente proyecto busca en una primera etapa, realizar la aplicación de la batería de riesgos psicosociales a una muestra de docentes en el área metropolitana de una de las ciudades capitales y efectuar entrevistas para conocer las variables de la organización del trabajo, que desde la perspectiva del profesor generan efectos negativos en el desempeño en los docentes de colegios y en la salud, para poder estructurar un primer modelo que logre identificar y determinar los factores de riesgo psicosocial presentes en el quehacer docente, luego, estructurar el conjunto de datos para la clasificación y el reconocimiento de patrones propios de esta actividad, para desarrollar adecuadamente el sistema de clasificación basado en Inteligencia Artificial que permita como herramienta de soporte al sistema de gestión de Seguridad y Salud en el Trabajo, actuar de manera más rápida en la intervención en los factores que están desencadenando riesgo psicosocial en docentes públicos.

Este documento se ha organizado para ser presentado de la siguiente manera: en primer lugar, se expone el marco teórico y referencial que da soporte a la identificación del objeto de estudio y al problema de investigación, y las bases conceptuales para su solución que compone la investigación. Seguidamente, se describe la metodología propuesta para la clasificación, luego, presenta los resultados y la discusión con base en los estándares de la metodología de la investigación científica. Finalmente, se exponen las conclusiones, las futuras líneas de investigación y las limitaciones.

II. Capítulo 1 – Planteamiento General

La presente tesis doctoral desarrolló en su totalidad tres capítulos: el marco teórico y referencial; diseño y aplicación de la metodología propuesta; experimentación y discusión de resultados y conclusiones y futuras líneas de investigación.

1.1. Vacío del conocimiento

En cuanto a los riesgos psicosociales que están presentes en la mayoría de los niveles en las organizaciones y que son considerados como riesgos emergentes (EU-OSHA, 2010a) se puede decir que:

Según el informe de la Comisión Mixta de la Organización Internacional del Trabajo, OIT, y la Organización Mundial de la Salud, OMS, (p.12), se define a los factores psicosociales en el trabajo dentro de la Seguridad y la Salud en el Trabajo como:

“Toda aquella interacción relacionada con el trabajo, el ambiente y la satisfacción en el trabajo relacionadas con la organización, la capacidad del trabajador, sus necesidades, su individualidad, su cultura y los factores externos al trabajo, que van ligadas a su percepción y experiencia, influye en las condiciones de salud, en la satisfacción con respecto a su trabajo y en el rendimiento del trabajador”.

Y también para la Organización Internacional del Trabajo, OIT, (1984):

Los factores positivos o negativos psicosociales relativos al trabajo comprenden aspectos del entorno de trabajo, del puesto de trabajo, de la organización, de las relaciones sociales entre trabajadores, de las relaciones del trabajador con su puesto de trabajo y el diseño de las tareas, las condiciones externas al trabajo y los aspectos propios de la personalidad del individuo (Sauter & Murphy, 1984, p. 1540).

Por otra parte, en Colombia los factores de riesgo psicosocial en el trabajo desde una mirada jurídica se han establecido en la Resolución 2646 de 2008 por el Ministerio de la Protección Social (Ministerio de la Protección Social, 2008) de la siguiente manera:

Establece y define las disposiciones para identificar, evaluar, prevenir, intervenir y monitorear constantemente las condiciones en las que los trabajadores estén expuestos a factores de riesgo psicosocial en el trabajo y el estudio de las patologías que se asocian a estrés ocupacional.

Los que se han definido de la siguiente manera:

Factores psicosociales. Son los aspectos pertenecientes al diario vivir del trabajador, relacionados con el ambiente organizacional, externo a la organización y propios del trabajador en relación con su familia, definidos como aspectos intralaborales, extralaborales e individuales que influyen en su desempeño.

“Y los factores de riesgo psicosocial intralaboral que se encuentran dentro de las empresas, son los siguientes:

Condiciones de la tarea, carga física y mental, jornadas de trabajo, beneficios que se reciben por parte de la empresa, condiciones ambientales de trabajo, jornada de trabajo, gestión de procesos, factores humanos, diseño de la organización del trabajo, interfaces humano-máquina y procesos de capacitaciones y formación”.

En el contexto mundial, los datos de los últimos años documentan un cambio en la salud (Seguridad y la Salud en el Trabajo) en las tendencias de la seguridad en Europa y en otras partes del mundo. Estos cambios han sido acompañados por el aumento de la prevalencia de los nuevos y emergentes tipos de riesgos a la salud y la seguridad (European agency for safety and health and work, 2011) de los trabajadores. Tal vez los nuevos retos en materia de seguridad y salud en el trabajo más ampliamente reconocidos son los riesgos psicosociales, los cuales están directamente relacionados con la expectativa de vida y con el producto interno bruto de un país (Dollard & Neser, 2013, p. 119).

Los riesgos psicosociales relacionados con el trabajo se refieren a aspectos del diseño y la gestión del trabajo, cognitivos, emocionales y comportamentales de los individuos y de sus contextos sociales y de organización que tienen el potencial de causar daño psicológico o físico (Sauter & Murphy, 1984).

Sobre la base de esta definición, todos los aspectos sociales y organizacionales del entorno de trabajo potencialmente pueden tener una influencia peligrosa sobre la salud de los empleados, entonces la gestión del riesgo psicosocial, se encuentra entre las responsabilidades de los empleadores para gestionar de manera preventiva y establecer los procedimientos para crear un excelente clima de seguridad psicosocial en los ambientes de trabajo (Dollard & Neser, 2013, p. 121).

Así mismo, Niedhammer & Chastang, (2013) mencionan lo siguiente: los factores psicosociales del trabajo están estrechamente asociados a las ausencias debido a enfermedad. Las demandas, la discriminación, la intimidación, la falta de identificación y

promoción de salud psicológica laboral y el desequilibrio en la vida laboral son factores de riesgo para la aparición de ausentismo por enfermedad para ambos sexos.

Se observa que, en Colombia, solo a partir del año 2004, a través del anterior Ministerio de la Protección Social y la Universidad de Antioquia, en el estudio denominado 'Violencia en el trabajo, formas y consecuencias de la violencia en el trabajo, Colombia 2004', se conoce que en los sectores de la educación y la salud aparecen los factores de riesgo psicosocial con una prevalencia nacional del 19.7% (Ministerio de la Protección Social, 2006).

Por otro lado, la validación a nivel internacional de instrumentos necesarios para identificar estrés laboral y condiciones psicosociales que afectan el ambiente de trabajo y que permiten a las gerencias adoptar medidas, van tomando mayor fuerza.

Por ejemplo, la Escuela de Negocios de Griffith de la Universidad de Griffith y la Escuela de Negocios de la Universidad de Queensland en Australia, desarrollaron y validaron un instrumento para evaluar la movilización de apoyo entre compañeros, el cual consta de 12 ítems que tiene una clasificación con referencia a las tres fuentes de apoyo (supervisor, compañeros de trabajo, la gente no perteneciente al trabajo), y evalúa la frecuencia con la que un empleado se ha acercado a cada una de esas fuentes para obtener cuatro funciones de apoyo (emocional, informativo, instrumental, evaluativa). El instrumento muestra evidencia de la fiabilidad de la dimensionalidad a través de una estructura factorial y la replicación de muestras, con validez convergente y discriminante a través de evaluaciones y resultados de la organización. Los valores obtenidos en la validación fueron una media de 1.88, una desviación estándar de 0.84 y un alpha de cronbach de 0.84 (Lawrence et al., 2015).

En otro estudio, llevado a cabo en el Research Institute on Addictions, de la Universidad Estatal de New York, Búfalo, Estados Unidos, se planteó desarrollar una definición conceptual y delinear las características deseables de una medida de la fatiga del trabajo,

examinar brevemente una serie de medidas previas de constructos relacionados con la fatiga de trabajo, y desarrollar y evaluar una nueva medida de la fatiga de trabajo. El trabajo denominado Inventario de Trabajo Fatiga Tridimensional (3D-WI FI) proporciona evaluaciones independientes y acordes de fatiga de trabajo físico, mental y emocional y obtuvo una consistencia interna de fiabilidad para las tres dimensiones de 0.94 para la fatiga física, de 0.95, para la fatiga mental y de 0.96, para trabajo emocional (Frone & Tidwell, 2015).

Se encontró también, que un grupo de investigadores del Departamento de Ciencias de la Salud de la Universidad de Groningen, Holanda, elaboró el cuestionario de desarrollo para medir el nuevo conjunto de capacidades para el trabajo, el cual muestra si un trabajador es capaz de apreciar y valorar su trabajo. Los resultados muestran la capacidad del cuestionario de trabajo para tomarlo como una medida de la empleabilidad sostenible basado en características propias de la relación ambiente de trabajo-individuo. La validez del constructo se realizó mediante pruebas de hipótesis, la puntuación para el conjunto de capacidades se realizó utilizando correlaciones de Pearson y análisis de regresión, las correlaciones de los siete aspectos de capacidad individuales del conjunto de capacidades estaban entre 0,555 y 0,710 (Abma et al., 2016)

Se debe mencionar que, a nivel mundial existen softwares para el registro de las variables que se miden en la identificación de riesgo psicosocial. Por ejemplo, el software psicsoft pro 3.0 de la empresa Psicopreven que contiene 3 metodologías para la identificación (F-psico 3.0 del Insht, Insl). F-psico es un procedimiento de evaluación psicosocial del Instituto de Seguridad e Higiene en el Trabajo de España, que se encuentra ya en la versión 3.1, el cual es una metodología de evaluación de factores psicosociales derivado de la Nota Técnica Preventiva Española (NTP 926 Factores psicosociales: metodología de evaluación; Cuixart, 2012) que informatiza los datos de dicha metodología.

En España, así mismo, se ha desarrollado un software para el manejo de la información generada al momento de evaluar las condiciones de trabajo y los riesgos psicosociales presentes en él. Se trata de la Metodología para la Evaluación y Prevención de los Riesgos Psicosociales, CoPsoQ-istas²¹, que deriva de la metodología CoPsoQ, que es un instrumento internacional para la investigación, la evaluación y la prevención de los riesgos psicosociales que tiene su origen en Dinamarca (Moncada et al., 2014).

En el año 2005 se llevó a cabo una tesis doctoral denominada “Diseño de un Sistema de Vigilancia Epidemiológica de Factores de Riesgo Psicosocial en el Trabajo”, cuyo principal objetivo era el diseño de una metodología para identificar factores de riesgo psicosocial en la población trabajadora de Colombia acorde a lo planteado por la legislación, ésta tesis basó su desarrollo en el modelo de Karasek de demanda, control y apoyo social y el modelo de Siegrist de desequilibrio, esfuerzo y recompensa (Karasek, 1979; Siegrist, 1996), que luego se usaría como base para el desarrollo del instrumento actual con el cual funciona la batería de riesgos psicosocial en Colombia (Villalobos, 2005).

En sus conclusiones se deja planteada la necesidad de avanzar en nuevas investigaciones en el campo de la identificación y prevención de los factores de riesgo psicosocial en Colombia, diseñando e implementando estrategias dirigidas a la promoción, prevención e intervención en los factores psicosociales y sus efectos en los trabajadores colombianos, creando herramientas e instrumentos que permitan el mejoramiento de las condiciones de salud.

Por consiguiente, en el año 2011, el Ministerio de la Protección Social diseñó una batería de instrumentos para la evaluación de factores de riesgo psicosocial para Colombia, adoptándose como base de evaluación el modelo propuesto por Villalobos (2005), el cual obtuvo un alpha de cronbach para el formulario de la Forma A de 0.954 y para la Forma B de 0.955, siguiendo lo indicado en la respectiva resolución dada por el ministerio.

Cabe anotar que la batería de instrumentos para la evaluación de factores de riesgo psicosocial de los trabajadores en Colombia que desarrolló el Ministerio de la Protección Social, cuenta con un aplicativo que compila los resultados (datos generales de la encuesta, puntaje transformado y niveles de riesgo) de todos los cuestionarios calificados en las encuestas cualitativas desarrollado bajo la plataforma Microsoft Excel 2007 que consolida la información con lo diligenciado de manera manual por los entrevistados.

Por consiguiente, se puede decir que los aplicativos informáticos disponibles hoy en día siguen una línea metodológica dada por un método que identifica riesgo psicosocial, pero que actúan como manejadores o gestionadores de la información a través de bases de datos que arrojan un resultado, pero no se enfocan en tratar el problema desde la inteligencia artificial a través de las técnicas inteligentes clasificando, reconociendo patrones presentes en los factores desencadenantes de los riesgos psicosociales y logrando predecir el grado de riesgo psicosocial que presentará un individuo.

En consecuencia, se observa que los instrumentos validados buscan siempre identificar las variables de satisfacción, de bienestar, de fatiga, de presión, tanto psicosociales como de las condiciones ambientales del trabajo que puedan ocasionar problemas de salud en el trabajador. También, podemos observar que los instrumentos desarrollados dejan abierta la posibilidad para realizar futuras mejoras y que la mayoría mencionan la necesidad de involucrar más variables que puedan aportar en los procesos de identificación, evaluación y control con tendencias hacia la prevención.

1.1.1. Situación problemática a abordar en la investigación

Se puede decir que muchos de los sistemas de trabajo y producción hoy en día se caracterizan por las altas demandas a los trabajadores, un fuerte énfasis en el trabajo que implica gran carga mental.

La ansiedad, la depresión y los trastornos psicosomáticos que se denominan de carácter psicosocial de corto plazo y los biológicos de largo plazo, como, infarto agudo, úlceras pépticas y dolores de espalda están directamente relacionados con las condiciones físicas de salud y de orden mental, atribuidas y relacionadas como efectos del estrés laboral, que se ha demostrado influyen en los sistemas osteomuscular, gastrointestinal, cardiovascular, respiratorio, endocrinológico e inmunitario (Vieco Gómez, & Abello Llanos, 2014).

De manera que estas problemáticas que son producto de las nuevas formas de trabajo que exigen del trabajador altas demandas laborales tanto físicas como mentales se traducen en presiones psicológicas que afectan al individuo y a las organizaciones (Mejía et al., 2009, p. 4) que deterioran la salud del trabajador acarreando múltiples problemas en todos los aspectos (físicos, psicológicos y emocionales) (García Castro & Muñoz Sánchez, 2013).

Los cambios que se presentan en los ambientes de trabajo junto a las altas demandas de trabajo que se le piden a los profesores generan entornos de trabajo estresantes, que producen la aparición de diversas afecciones que afectan la salud mental y que directamente influyen en la pérdida de interés en la labor desempeñada, con efectos negativos en todos los procesos, tanto de relación con el estudiante, como de tipo administrativo con sus compañeros y jefes (Olaya Arévalo, 2015).

En Colombia, algunos estudios han dejado ver la magnitud de los riesgos existentes en la profesión docente de educación básica primaria y secundaria, por ejemplo, Diazgranados et al. (2006), realizaron entrevistas y grupos focales con maestros y rectores, creando una encuesta de donde obtuvieron información cualitativa que se aplicó a 562 docentes de 25 colegios distritales en Bogotá, obteniendo resultados donde se identifican que características inciden negativamente en los maestros, y se encontrando diferencias entre la salud física, emocional y psicológica.

Los resultados muestran que en cuanto a la salud física los problemas más comunes son trastornos vocales, respiratorios, afecciones osteomusculares y fatiga, en el caso psicológico, los más comunes son dificultad para concentrarse y pérdida de la memoria y en la parte de desempeño laboral, se encontró desmotivación frente al trabajo, conflictos interpersonales, ausentismo e insatisfacción laboral.

En el año 2011, un estudio llevado a cabo en la ciudad de Pereira, denominado Factores de riesgo psicosocial que afectan el ejercicio pedagógico de los docentes, que seleccionó 100 docentes de 11 colegios, señalaba lo siguiente en cuanto a alteraciones físicas asociadas a situaciones estresantes (Restrepo, 2011):

Desespero en su actividad académicas (68%), irritarse con facilidad en el trabajo (60%), frecuentes dolores de cabeza (60%), migraña y cefalea tensional (60%), dolor en cuello y brazos (60%), dolor en hombros (52%), sueño irregular y malestar en la noche (52%) y síntomas de depresión (36%).

Entonces, las características del trabajo docente enmarcan dentro de su trabajo el tener que redefinir las formas en que desarrolla su trabajo y modificar las alternativas que se deben seguir para poder prevenir enfermedades laborales. Los principios de higiene y seguridad, así como el de organización del trabajo, no son suficientes para definir la tarea docente. En su estudio Restrepo (2011, p. 77) menciona lo siguiente: “Los docentes que son sometidos a todo tipo de presiones, en un medio de conflictos y donde encuentran carencias y falta de apoyo, presentan sintomatologías severas de tipo psíquico que dificulta el que logren transmitir adecuadamente el conocimiento a sus alumnos”.

En Bogotá, también, el estudio llevado a cabo por (Espinoza & Villalobos, 2015, p. 14), observó que la sintomatología por estrés muy alto y alto equivale al 62%, asociados al liderazgo y a las relaciones sociales en el trabajo, al control sobre el trabajo, a las demandas del trabajo y las recompensas, factores a tener en cuenta para diseñar y desarrollar estrategias de intervención basados en programas que se puedan implementar.

La conclusión de esta investigación es la de crear urgentemente programas de vigilancia epidemiológica junto a programas de apoyo y afrontamiento que reduzcan el impacto en la salud mental de compañeros y directivos, además de nuevos estudios que permitan comparar resultados para generar políticas de promoción y prevención que se puedan aplicar al sector.

De manera que, las investigaciones en Colombia que dejan ver la relación existente entre el bienestar mental y las condiciones psicosociales del entorno laboral, son las principales razones que motivan a la investigación y el desarrollo de nuevas herramientas que permitan la intervención de los factores negativos de riesgo psicosocial relacionados con el trabajo, entendiendo que estos se transforman en coadyuvantes en el deterioro de la salud y la capacidad laboral, reduciendo las habilidades, destrezas y potencialidades de los trabajadores (Villalobos F, 2004).

Es de notar que el sector educativo es uno de los sectores de la economía más expuesto a enfermedades de factor psicosocial (Ministerio de la Protección Social 2006). Dentro de la segunda encuesta nacional de condiciones de seguridad y salud en el trabajo en el sistema general colombiano de riesgos laborales (Ministerio del Trabajo & Organización Iberoamericana de Seguridad Social, OISS, 2013), 14.076 centros de trabajo pertenecían a la educación dentro de la encuesta y se halló que el 21.11% de la población presentó problemas de tensión en el trabajo y solo el 23.27% sintieron que con frecuencia pudieron solucionar sus problemas, el 8.76% declaró haberse sentido triste o deprimido, el 12.37% declararon haber perdido la confianza en sí mismos y el 5.50% refería nunca sentirse feliz. Estos hallazgos muestran la urgente necesidad de intervenir y realizar acciones para promover la salud y prevenir los trastornos mentales en los trabajadores, incluidos los docentes.

Hay que tener en cuenta que a pesar que parece que la labor docente posee un riesgo bajo, realmente los riesgos psicosociales que se encuentran en esta labor son muchos, y esto se puede observar en lo reiterado de las consultas médicas debido a problemas osteomusculares, como dolores de cuello y espalda, problemas digestivos, como gastritis,

problemas cardiovasculares, como hipertensión, los que con frecuencia se deben a consecuencia del estrés, lo que indica el crecimiento negativo de los riesgos psicosociales y afecta la salud de los maestros, como ya lo han dejado ver las investigaciones que se han llevado a cabo. (Desespero en su actividad académicas (68%), irritarse con facilidad en el trabajo (60%), frecuentes dolores de cabeza (60%), migraña y cefalea tensional (60%), dolor en cuello y brazos (60%), dolor en hombros (52%), sueño irregular y malestar en la noche (52%) y síntomas de depresión (36%), (Restrepo, 2011) (**Figura 1- 1**).

Kortum et al., (2010) relacionan que la monitorización y vigilancia de los riesgos psicosociales son prioridad en todos los países en desarrollo, dejando indicada la necesidad de incluir cuestiones relativas a los riesgos psicosociales en las políticas nacionales y en los marcos legislativos de salud y seguridad en el trabajo, en donde Colombia ha hecho grandes esfuerzos en los últimos años a través de la Ley 1562 de 2012 (Ministerio del Trabajo, 2012), el Plan Decenal de Salud Pública 2012-2021 y el Plan Nacional de Seguridad y Salud en el Trabajo 2013-2021, en el que **se detecta la necesidad de diseñar, desarrollar e implementar herramientas y estrategias que logren identificar, prevenir y controlar los factores de riesgo psicosocial en todos los sectores productivos del país**, con el fin de mejorar las condiciones de salud de los trabajadores, que faciliten el realizar el seguimiento a las condiciones diagnosticadas y mostrando la evolución de las medidas tomadas a través de técnicas estadísticas que arrojen valores que sean indicadores de la gestión y permitan fortalecer los programas de promoción de la salud y prevención de la enfermedad.

1.2.1. Relación factores de riesgo psicosocial e inteligencia artificial

Posada-Quintero et al., (2020) desarrollaron la medición de ciertos problemas relacionados con el Burnout en Colombia, como el trabajo de horas extras, falta de comunicación con los jefes, salario inadecuado y falta de apoyo por parte de los jefes frente a los padres de familia. El modelo de aproximación utilizó técnicas de machine learning como máquinas de

soporte vectorial y arboles de decisión para facilitar la predicción del síndrome de burnout en docentes de colegios colombianos. El uso de técnicas lineales ha demostrado ser importante como herramienta para la predicción de condiciones psicosociales que afectan la salud del trabajador.

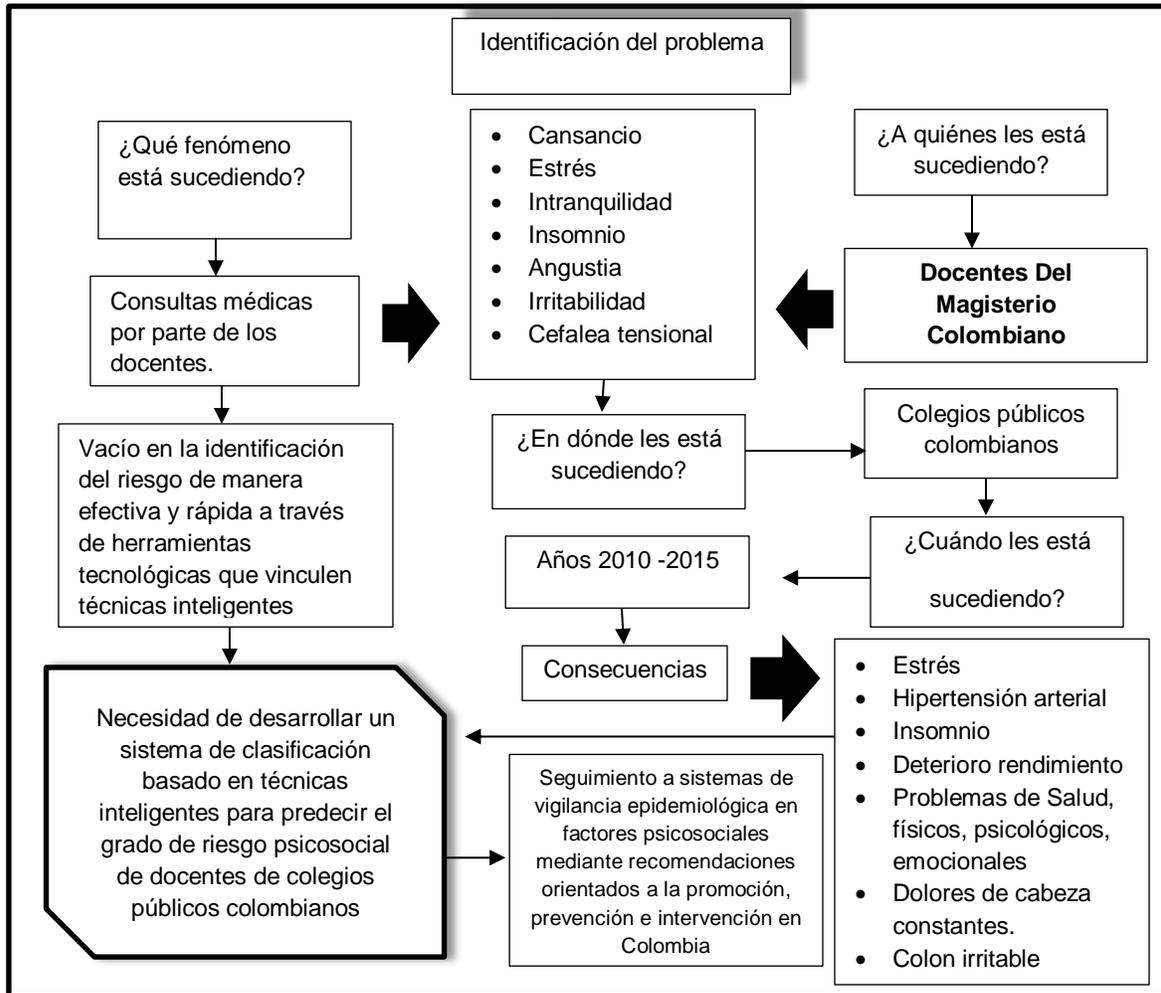
Como lo analiza Czaja & Nair, (2006), la relación que tienen los factores de riesgo psicosocial con las técnicas inteligentes o técnicas de inteligencia artificial está estrechamente ligada a la concepción que desde la ingeniería de factores humanos y diseño de sistemas se le puede dar. Las organizaciones de trabajo y los entornos sociales han cambiado enormemente, y estos cambios continuarán conforme avanza la tecnología y los patrones sociales evolucionan. Estos cambios continuarán a la vez que surgen nuevas tecnologías y que se continúan agilizando los procesos en la economía del sector servicios.

El concepto de sistema fue inicialmente una filosofía asociada con pensadores como Hegel, que reconoció que el todo es más que la suma de sus partes. También era un concepto fundamental entre los psicólogos de la Gestalt, que reconocieron la importancia de "objetualidad" o la totalidad de la percepción humana. El enfoque sistémico, que evolucionó a partir de los sistemas de pensamiento, fue desarrollado inicialmente en las ciencias biológicas y refinado por los ingenieros de comunicación en la década de 1940. La adopción de este enfoque se vio reforzada durante la Segunda Guerra Mundial, cuando se reconoció que los sistemas militares se estaban volviendo demasiado complejos para que los seres humanos los pudieran operar con éxito. Este descubrimiento dio lugar a la aparición del campo de la ingeniería de los factores humanos y su énfasis en los sistemas hombre-máquina.

En este sentido, para diseñar los sistemas de trabajo de hoy en día de manera efectiva, tenemos que aplicar los conocimientos respecto a las capacidades de procesamiento de información humano en el proceso de diseño. La necesidad de este tipo de conocimiento ha creado un mayor énfasis en las cuestiones relacionadas con la cognición humana

dentro del campo de los factores humanos y ha dado lugar a la aparición de la ingeniería cognitiva.

Figura 1- 1. Identificación de síntomas de educación básica primaria y secundaria en colegios públicos de Colombia. Adaptado de: Espinoza (2015); García (2010); Diazgranados (2006), Restrepo (2011).



Fuente: elaboración propia.

Ésta se centra en la complejidad, el pensamiento cognitivo y en los conocimientos relacionados con la actuación humana, ya sean llevadas a cabo por los seres humanos o por los agentes de máquina. Está estrechamente alineado con el campo de la ciencia cognitiva y la inteligencia artificial. Con énfasis en el trabajo en equipo, el concepto de la cognición en equipo ha surgido haciendo referencia a la interacción entre los procesos

cognitivos intraindividual e interindividual y aplicando las herramientas conceptuales de la ciencia cognitiva para un equipo o grupo en comparación con el individuo.

Lo que quiere decir que: *“los procesos organizacionales deben garantizar un diseño de los ambientes de trabajo de manera que las condiciones macrocognitivas presentes en los lugares de trabajo conlleven un análisis profundo y un diseño desde el punto de vista macroergonómico, lo que vincula el diseño de puestos de trabajo y el diseño de las tareas en los entornos garantizando ambientes de trabajo saludables con el uso de todas las técnicas y tecnologías al alcance de quienes diseñan los lugares y los procesos de trabajo”*.

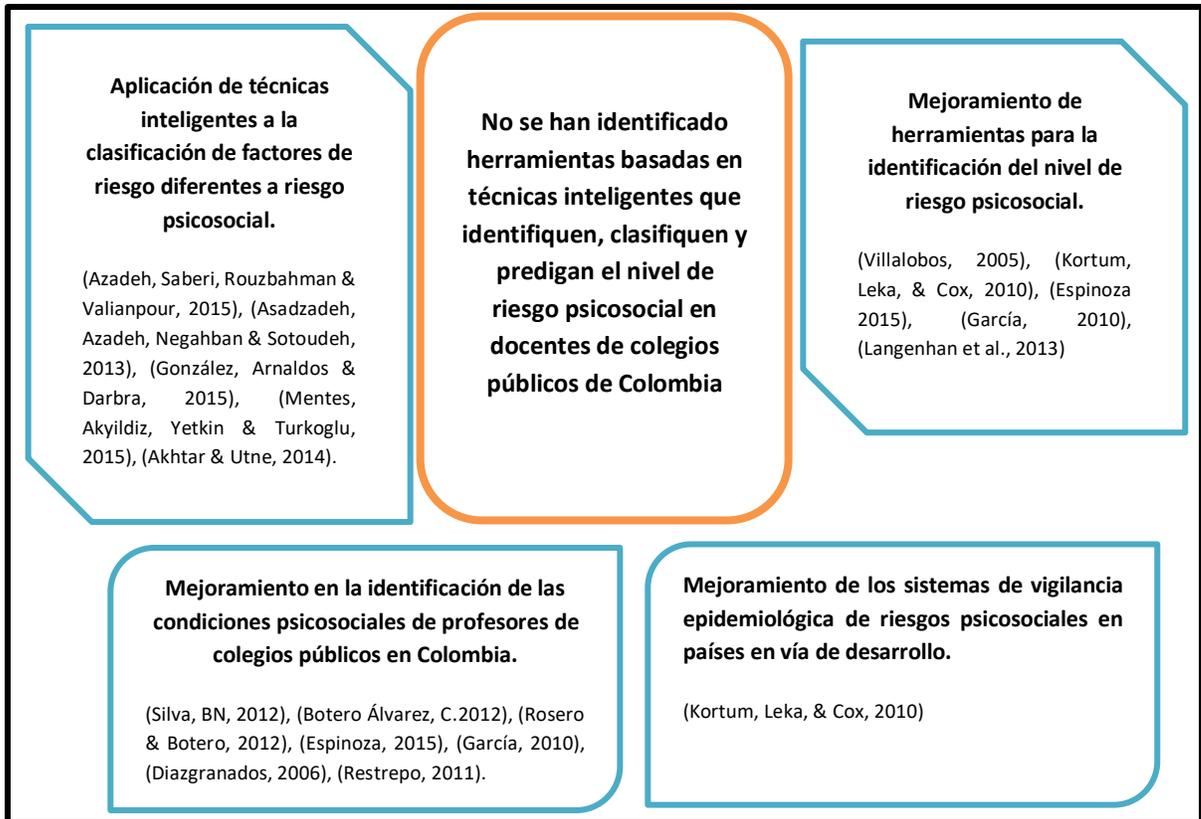
En conclusión, de la revisión de literatura, es posible establecer que se deben desarrollar metodologías que busquen fortalecer la prevención de los riesgos psicosociales a través de la predicción en Colombia. Esto se puede lograr a través de técnicas inteligentes que faciliten el proceso automatizado, en diferentes etapas, de todo el proceso de vigilancia epidemiológica, tanto de identificación, clasificación y predicción. La **Figura 1- 2** esquematiza los aspectos principales derivados de la revisión de la literatura científica y técnica, sobre los cuales se sustenta el vacío de conocimiento identificado en el presente proyecto de investigación.

Por consiguiente, se identifica entonces:

La necesidad del diseño de un sistema de clasificación basado en técnicas inteligentes que clasifique y prediga el grado de riesgo psicosocial con aplicación en profesores de colegios públicos en Colombia, que contribuya como una herramienta tecnológica de soporte, al mejoramiento del diagnóstico por parte del psicólogo laboral de los factores de riesgo psicosociales presentes en los ámbitos educativos, que mejore y reduzca los tiempos de identificación y diagnóstico y optimice las condiciones de calidad y las acciones de promoción de la salud, prevención de la enfermedad laboral y permita la intervención de los factores psicosociales y sus efectos asociados en los docentes, todo ligado a la innovación.

Basado en lo anterior, se puede realizar la sistematización del problema de investigación.

Figura 1- 2. Vacío del conocimiento.



Fuente: Elaboración propia.

1.2. Sistematización del Problema

De acuerdo con los planteamientos anteriores, y el vacío de conocimiento detectado en la revisión sistemática de la literatura, la pregunta de investigación que reúne los aspectos claves mencionados y que guiará el desarrollo del presente proyecto de investigación, es la siguiente:

¿Cómo desarrollar un sistema de clasificación basado en técnicas inteligentes para la predicción del grado de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia?

Subpreguntas de investigación

- Conceptualmente, ¿cuáles son los sistemas de clasificación basados en técnicas inteligentes que permiten desarrollar metodologías para identificar, clasificar y predecir el nivel de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia?
- ¿Qué variables independientes (nuevas y existentes) más efectivas se pueden incluir en un sistema de clasificación, con el fin de lograr un mayor nivel de automatización en el proceso de clasificación y predicción del grado de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia?
- ¿Cómo se debe diseñar y desarrollar un sistema de clasificación basado en técnicas inteligentes que permita identificar, clasificar y predecir el nivel de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia?
- ¿Cómo validar el sistema de clasificación propuesto con el fin de establecer su efectividad en comparación con los métodos empleados por los especialistas?

1.3. Justificación

En muchos países de Europa también se han incrementado los riesgos psicosociales (incluyendo el acoso e intimidación), y el estrés relacionado con el trabajo como

consecuencia de la acumulación, con pruebas sobre su prevalencia e impacto. En algunos casos, por ejemplo, en Italia, éstas han dado lugar a cambios en la legislación con el aumento de la participación de las organizaciones a todo nivel (Houtman et al., 2020; Iavicoli et al., 2014). Además, en algunos países de la UE, la legislación es aún más específica que la ley de la UE y hace referencia directa al estrés relacionado con el trabajo, la intimidación y el acoso o riesgos psicosociales, aunque en muy pocos países las enfermedades asociadas con el estrés son incluidas en las listas oficiales de enfermedades profesionales.

En consecuencia, la idoneidad de los métodos y herramientas disponibles para la identificación de riesgos psicosociales se considera como deficiente con respecto a las capacidades existentes en el contexto de los cambios socioeconómicos de la competitividad a nivel mundial (Bauer & Hämmig, 2014). El Instituto Sindical Europeo (Etui, 2013) organizó el seminario sindical europeo de riesgos psicosociales en Bilbao, con el apoyo de la EU-OSHA, en donde también se indicó la necesidad de mejorar los instrumentos para la identificación de este tipo de riesgos. Surgen entonces dificultades en la comprensión sobre el "tradicional" punto de vista de la salud y la seguridad. Sobre la base de la gestión de riesgos, las empresas tratan el "riesgo" y "la gestión del riesgo" de forma rutinaria en áreas tales como finanzas, estrategia y operaciones (entre otros) visualizando sólo la parte física y económica de los procesos (Langenhan et al., 2013, p. 92).

Por ello, se debe proteger a través de prácticas responsables a nivel de política, de negocios y de responsabilidad social corporativa a los empleados, con la implementación y el monitoreo de instrumentos y políticas que conlleven a la reducción del riesgo, incluidos los riesgos psicosociales (Bauer & Hämmig, 2014, p. 237). Asimismo, se debe entrenar en conocimientos y desarrollo de competencias a los actores clave en esta área, tales como inspectores de trabajo y a las personas involucradas con los Sistemas de Gestión de Seguridad y Salud en el trabajo.

Como lo sugieren Rosero & Álvarez, (2012), “*existe la necesidad en el país de desarrollar instrumentos que permitan la identificación y evaluación de los factores psicosociales que afectan la salud y el desempeño en docentes de Colombia*”. Contrasta este direccionamiento con lo señalado por Cox et al., (2006) en el documento denominado ‘Intervenciones organizativas para el estrés laboral. Un enfoque de gestión de riesgos’, se señala que los métodos de evaluación del riesgo que produce estrés en el trabajo son escasos y la mayoría se limita solo a dar recomendaciones didácticas que siguen un diagrama de flujo. Plantea que entonces se hace necesario recurrir a la lógica creativa que permita el diseño de herramientas de identificación que realicen una trazabilidad y la identificación temprana sobre los procesos psicosociales.

Es importante, entonces, corroborar observaciones sobre los factores de riesgo psicosocial en el sector educativo público en Colombia, con instrumentos nacionales y en escuelas y colegios del país, en especial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia, los cuales pueden ser vulnerables por condiciones de salud que afectan su desempeño y el desarrollo normal de la vida.

Según lo planteado por Posada-Quintero et al., (2020), el docente en medio de su labor padece de afecciones a causa del entorno psicosocial que le rodea, es vital entonces, procurar su bienestar y satisfacción en pro de su labor y salud. Por lo tanto, se hace necesario adoptar nuevas formas de para la identificación de riesgos psicosociales que permitan la promoción y la prevención de la salud en los ambientes de trabajo de cada docente.

En otras palabras, el bienestar físico y mental de los docentes está estrechamente relacionado con las características organizacionales de sus empleos, es necesario crear estrategias de apoyo individual y grupal que fortalezca su condición psicológica y permita disminuir o controlar los síntomas negativos que inciden o prevalecen sobre su salud.

Entonces, se debe tener en cuenta que en Colombia el Riesgo psicosocial está clasificado como enfermedad laboral (Agente Psicosocial) que se encuentra definido en la tabla de enfermedades laborales según Decreto 1477, 2014 por el cual se expide la tabla de enfermedades laborales (Ministerio del Trabajo, 2014, p. 26).

Por lo tanto, se considera pertinente investigar el fenómeno y sus efectos en docentes del magisterio colombiano, la Ley 1562, 2012, define que la Seguridad y la Salud en el Trabajo es una disciplina que es la encargada de la prevención de las lesiones y enfermedades relativas al trabajo y que se encarga del mejoramiento de las condiciones ambientales y del entorno del trabajo para promover el bienestar físico, mental y social de todas las personas en los ambientes de trabajo.

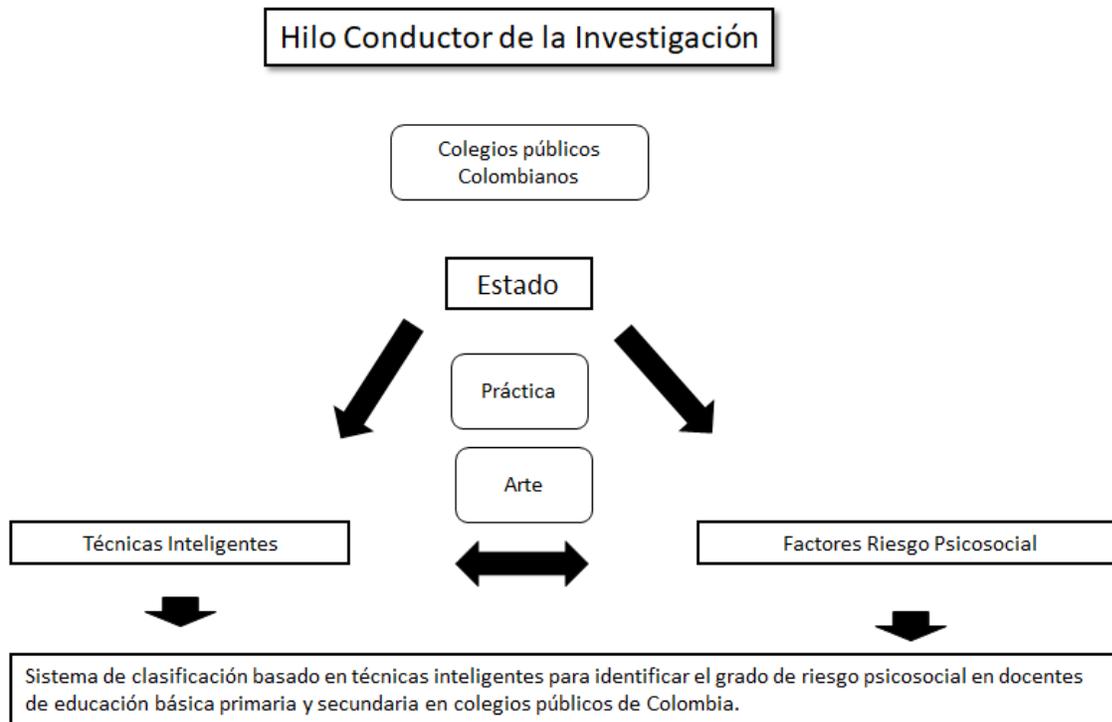
Acorde a lo planteado en esta Ley, se deben establecer programas y servicios que propendan por la promoción de la salud y la prevención y control de los riesgos laborales. Y como lo plantea el Decreto 1072, 2015, Decreto Único Reglamentario del Sector Trabajo, es necesario implementar metodologías adicionales que complementen la evaluación de peligros de tipo psicosocial. Además, el Plan Decenal de Salud Pública 2012-2021 prioriza la identificación de los riesgos ocupacionales y los efectos que tengan sobre la salud de los trabajadores de los riesgos ergonómicos y psicosociales.

También, como lo señala y promueve el Ministerio del Trabajo en la Resolución 06045, 2014, se adopta el Plan Nacional de Seguridad y Salud en el Trabajo 2013-2021, que fortalece la Seguridad y la Salud en el Trabajo y la prevención de riesgos laborales para el sector educación, sensibilizando a los directivos de diversos niveles para promocionar y crear estrategias que divulguen los conceptos de salud en el trabajo. El Ministerio de Educación Nacional, a través del Decreto 1075, 2015, Único Reglamentario del Sector Educación, busca y propende por el bienestar psicosocial de los docentes de educación básica primaria y secundaria en colegios públicos de Colombia, enfatizando en la creación de los perfiles de riesgo laboral de cada docente relativo a factores de riesgo psicosocial, formulando, diseñando e implementando estrategias que busquen intervenir y reducir los riesgos presentes en los entornos educativos.

1.4. Diseño general de la investigación

La estrategia que siguió el autor para identificar el “estado del arte y de la práctica” para la construcción del marco teórico - referencial de esta investigación (hilo conductor) se sustenta en el esquema que se expone en la **Figura 1- 3**.

Figura 1- 3. Hilo conductor de la investigación.



Fuente: elaboración propia.

Por lo cual se procedió a buscar en la literatura científica los temas relevantes en sistemas de inteligencia artificial computacional, específicamente en lógica difusa, fuzzy sets, árboles de decisión, redes neuronales, redes bayesianas, máquinas de soporte vectorial, algoritmos aleatorios, algoritmos genéticos y otras técnicas que puedan ser aplicadas a

procesos de Seguridad y Salud en el Trabajo, para la identificación, clasificación y predicción del nivel de riesgo psicosocial. Los hallazgos se abordarán de forma general en el proceso organizacional de prestación del servicio en docentes de educación básica primaria y secundaria en colegios públicos de Colombia.

La búsqueda sistemática se realizó en las plataformas:

ISI Web of Science, Scopus y Dialnet.

Bases de datos científicas: Science Direct, Springer Journal, Taylor & Francis y Wiley Ebscohost, Cochrane, Emerald, Medline, Pubmed, Redalyc, Scielo. La estrategia sistemática de búsqueda y las ecuaciones de búsqueda se pueden consultar en el anexo D.

1.5. Hipótesis general de la investigación

Un sistema de clasificación basado en técnicas inteligentes soportado en las variables psicosociales es capaz de medir el grado de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia, con una efectividad igual o superior al 95% comparado con las técnicas de medición psicosocial vigentes en Colombia.

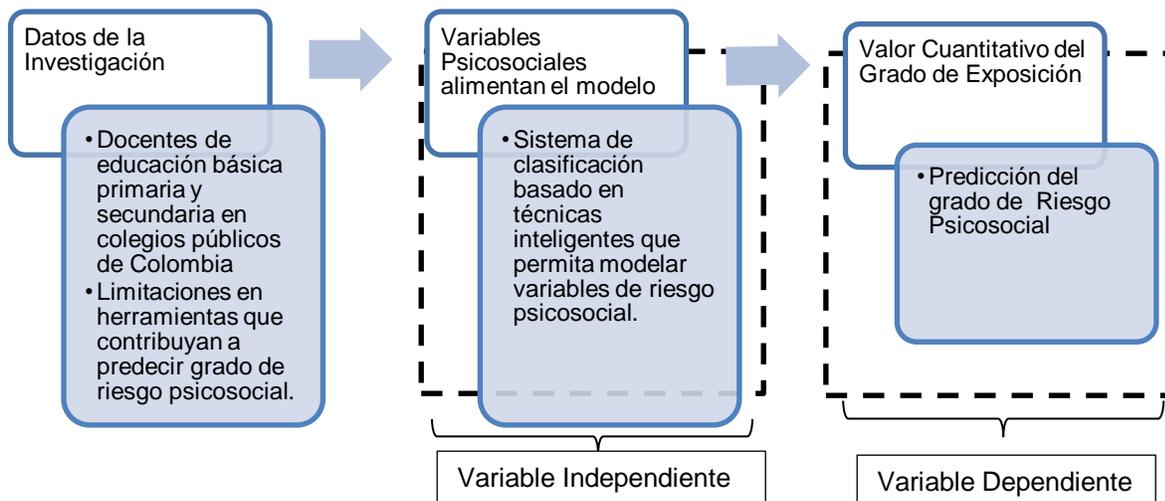
1.5.1. Identificación y definición conceptual y operacional de las variables

Sistema de clasificación Inteligente: Programa computacional que utiliza variables y características ya conocidas de alguna situación con las cuales interpreta

probabilísticamente sucesos o hechos que dan un resultado, el cuál es un objetivo programado previamente.

Grado de riesgo psicosocial: Conjunto de condiciones que afectan al individuo de forma negativa en su entorno laboral, generando efectos perjudiciales en docentes de educación básica primaria y secundaria en colegios públicos de Colombia, que pueden producir una condición desfavorable en su salud (**Figura 1- 4**).

Figura 1- 4. Operacionalización de variables.



Fuente: Elaboración propia.

1.5.2. Comprobación de la hipótesis de investigación

La aplicación y validación del sistema de clasificación propuesto con el fin de establecer su efectividad en comparación con los métodos tradicionalmente empleados por los especialistas en docentes de educación básica primaria y secundaria en colegios públicos de Colombia permite identificar el grado de riesgo psicosocial igual o superior al 95%.

1.5.3. Tipo de investigación

Esta investigación, según el tipo de investigación, es exploratoria, pues de manera científica se abordará la identificación de factores de riesgo psicosocial a través de técnicas computacionales inteligentes; es correlacional, pues estudia la relación entre variables dependientes e independientes, según la naturaleza de los datos de metodología cuantitativa por el empleo estadístico que se hará de los datos y por la manipulación de las variables; es de tipo experimental, por la necesidad de manipular una o más variables independientes para conseguir los resultados esperados.

1.6. Objetivo General

Diseñar y desarrollar un sistema de clasificación basado en técnicas inteligentes para predecir el nivel de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia con una efectividad igual o superior a las técnicas actualmente existentes.

1.7. Objetivos Específicos

- Analizar los sistemas de clasificación basados en técnicas inteligentes y no inteligentes que sirvan para diseñar modelos para la identificación, evaluación y seguimiento de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia.
- Definir las variables independientes (nuevas y existentes) en la identificación de los riesgos psicosociales y seleccionar las variables más efectivas mediante técnicas

informáticas y/o estadísticas, con el fin de lograr un mayor nivel de automatización en el proceso de predicción de este riesgo.

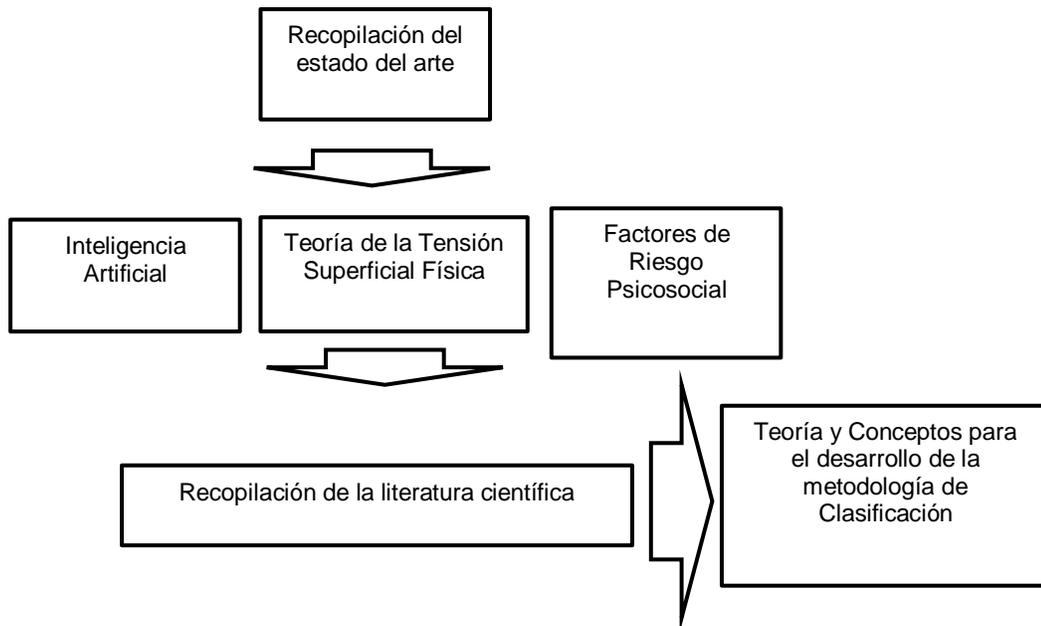
- Desarrollar un sistema de clasificación basado en técnicas inteligentes que permita identificar, clasificar y predecir el nivel de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia.

- Validar el sistema de clasificación propuesto con el fin de establecer su efectividad en comparación con los métodos tradicionalmente empleados por los especialistas aplicados en docentes de educación básica primaria y secundaria en colegios públicos de Colombia.

III. Capítulo 2 - Marco teórico y referencial de la investigación

El marco teórico y referencial de la investigación aborda tres tópicos: primero, la inteligencia artificial como principal técnica para resolver el problema de clasificación del grado de riesgo psicosocial, segundo, la teoría de la tensión superficial física, que sirve como elemento innovador dentro del modelo de clasificación y por último, el riesgo psicosocial como elemento asociado a la productividad laboral en los entornos laborales, principalmente dentro del sector educativo de colegios públicos. Con lo anterior, se establecen las bases teóricas y conceptuales del desarrollo de la metodología propuesta (Figura 2-1).

Figura 2 - 1. Esquema del marco teórico y referencial de la investigación.



Fuente: Elaboración propia.

2.1. Inteligencia Artificial

La idea de tener máquinas inteligentes surge en el siglo XVII, cuando John Locke planteó si los equipos y máquinas podrían tener inteligencia. Alan Turing, definió que para que una máquina fuera inteligente debería realizar tareas en contextos específicos de manera similar a un hombre, la cual fue programada previamente (Santos, 2011, p. 285).

Turing, fue pionero en definir la computabilidad y abrió el camino para el nuevo campo que hoy en día se conoce como las ciencias de la computación y las ciencias cognitivas (Teuscher, 2004), la abstracción del proceso creativo ha sido un estudio central para los enfoques cognitivos y uno de los terrenos más fértiles para la cosificación matemática de ideas en la investigación de la Inteligencia Artificial. Desde ese punto de vista, la Inteligencia Artificial está obligada a imitar los esfuerzos creativos humanos al reproducir la abstracción y las tácticas que entran en juego en nuestro cerebro. En ese sentido, los enfoques cognitivos se han centrado principalmente en procesos fundamentales que pueden traducirse en términos computacionales. Esto implica la medida en que se organizan y acceden al conocimiento y la información, a través de diferentes tipos de sistemas de memoria, y los correspondientes procesos de recuperación y análisis de diversas fuentes de información que permiten clasificar adecuadamente situaciones de la vida real (Esling & Devis, 2020).

El problema que conlleva dar la definición sobre lo que es un sistema inteligente consiste en las diversas formas de inteligencia. Para que un organismo se pueda definir como inteligente, requiere como mínimo emular la habilidad de percibir, adaptarse, tomar decisiones y realizar acciones de control.

En resumen, un sistema inteligente debe lograr apropiadamente en un entorno bajo incertidumbre, conseguir acciones apropiadas que aumenten las probabilidades de éxito, el cual debe considerar cumplir con los sub-objetivos planteados que logren la consecución del objetivo principal del sistema (Santos, 2011).

Existen varias metodologías inteligentes, por ejemplo, la lógica difusa, algoritmos de búsqueda, redes neuronales artificiales, redes bayesianas, programación evolutiva, algoritmos genéticos y optimización heurística (Benítez et al., 2014; Santos, 2011).

Refiere Tettamanzi & Tomassini, (2001) que “desarrollos recientes en áreas como las redes probabilísticas, la lógica difusa, los algoritmos evolutivos y las redes neuronales artificiales son ingredientes esenciales del soft computing, los cuales son todos bioinspirados y pueden ser fácilmente combinados sinérgicamente”.

El utilizar esta sinergia, es lo que ha resultado muy útil en diferentes campos donde se ha aplicado, muchos de los métodos de inteligencia artificial se han utilizado en sectores industriales, académicos y de investigación con resultados óptimos para el objetivo planteado gracias a la combinación de ellos.

(Benítez et al., 2014; Duda et al., 2001) exponen:

El principal objetivo de la aplicación del aprendizaje automático (*machine learning*) utilizado para clasificación es el de lograr una aproximación (*modelo*) a partir de una serie de atributos dados (*Datos*) desordenados para lograr hallar un comportamiento del cual se pueda inducir una serie de resultados que tiendan a generalizarse para poder entenderlos (*Clase*).

El conjunto de entrenamiento $S = \{(x_1, y_1), \dots, (x_n, y_n)\}$ contiene n ejemplos de entrenamiento, que corresponden a pares (x, y) donde $x \in X$ e $y = f(x)$. La componente x de cada ejemplo es un vector $x = (x_1, \dots, x_n)$ de atributos con valores continuos o discretos que describen la información relevante o propiedades del ejemplo. Los valores del espacio de salida Y asociados a cada ejemplo son las clases del problema. Así, cada ejemplo de entrenamiento queda totalmente caracterizado por un conjunto de pares atributo–valor y una etiqueta de clase (Benítez et al., 2014; Duda et al., 2001).

La teoría que trata del aprendizaje automático para clasificación (Aprendizaje supervisado) viene de la teoría del aprendizaje estadístico (Statistical learning theory) (Fernandes de Mello & Antonelli, 2018; Vapnik, 1995). Para la construcción de modelos se hace necesario recurrir al sesgo y la varianza, donde el área de búsqueda se relaciona con el sesgo y el grado de complejidad del modelo se identifica a través de la varianza, que mide los valores dispersos (Belkin et al., 2019; Rosset & Tibshirani, 2020) y se utilizan en diferentes algoritmos supervisados, incluidas las redes neuronales (Neal et al., 2019).

Como se puede leer e inferir de (Alpaydin, 2020; Rosset & Tibshirani, 2020):

Los algoritmos están basados en reglas o probabilidades las cuales se utilizan para adquirir los modelos o también llamadas reglas de clasificación de los datos ejemplos, con los que al combinarlos pueden lograr ser más eficientes.

2.2. Aprendizaje Supervisado

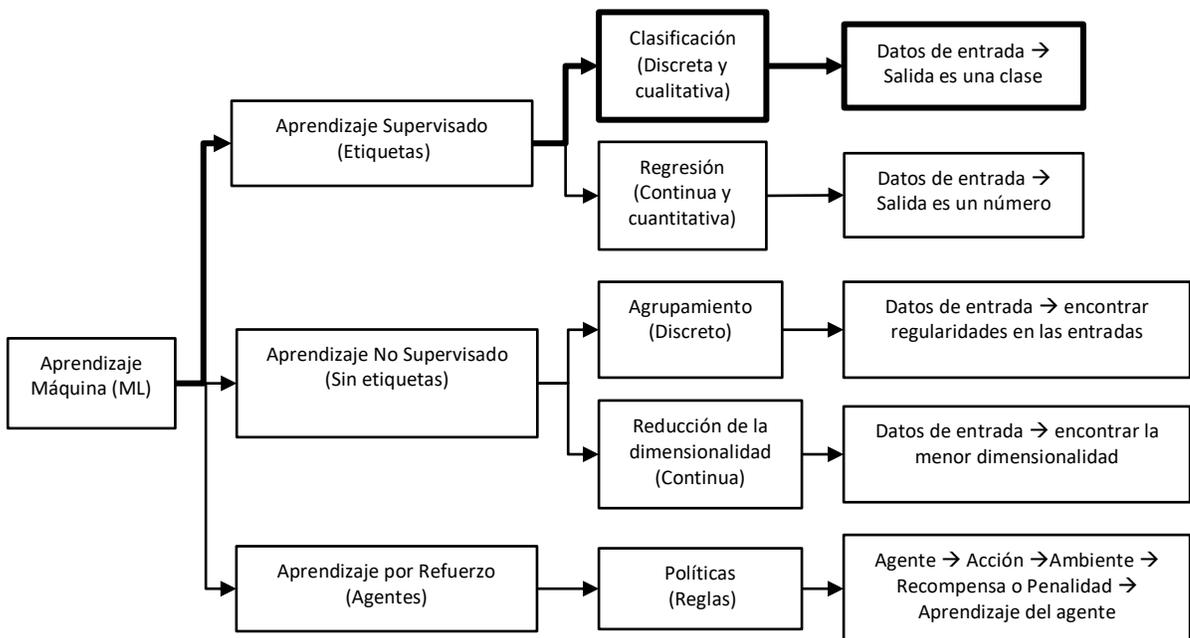
El aprendizaje supervisado (**Figura 2- 2**), es una técnica que logra hallar una función a partir de una serie de datos de entrada, llamados de entrenamiento, que son vectores, y unos datos los cuales se deben hallar, denominados resultados (que son etiquetas de clase o valores numéricos) (Rashidi et al., 2019). El objetivo es lograr que la función logre clasificar o predecir el mayor número de datos correctos a partir de una serie de datos de ejemplo (datos de entrenamiento) con datos no vistos previamente (Murdoch et al., 2019).

2.2.1. Clasificación

La clasificación es una técnica de aprendizaje supervisado que toma muestras de datos etiquetados previamente definidos y genera un modelo (llamado clasificador) que clasifica

nuevas muestras de datos en diferentes grupos o clases predefinidos (Belkin et al., 2019; Han et al., 2012). En los procesos de minería de datos y de aprendizaje automático la clasificación ha sido estudiada y aplicada ampliamente (Alpaydin, 2020; Holte, 1993; Kubat, 2017; Rashidi et al., 2019).

Figura 2- 2. Jerarquía de diferentes algoritmos de aprendizaje automático que incluyen técnicas de aprendizaje supervisadas versus no supervisadas versus reforzadas.



Fuente: Adaptado de (Rashidi et al., 2019).

Los algoritmos de aprendizaje de máquina (Machine learning) de tipo clasificación, obedecen a la técnica de aprendizaje supervisado, la cual puede manejar datos de tipo cuantitativo o cualitativo, y su resultado es generalmente llamado clase.

2.2.2 Métodos de clasificadores lineales

Un clasificador lineal es un hiperplano en un espacio *n-dimensional* de atributos que puede ser representado por un vector de pesos w y un umbral b que está relacionado con la distancia del hiperplano al origen: $h(x) = \text{signo}(\langle w, x \rangle + b)$ (donde $\langle w, x \rangle$ representa el producto escalar). Cada componente del vector de pesos se corresponde con uno de los atributos (Vapnik, 1995; Widrow & Hoff, 1960).

2.2.3 Redes Neuronales Artificiales

McCulloch y Pitts modelaron la primer estructura y funcionamiento simplificado de las neuronas cerebrales, describiéndolas como dispositivos con n entradas, una única salida y solo dos estados posibles: activa o inactiva (McCulloch & Pitts, 1943).

Las Redes de Neuronas Artificiales (RNA), tienen la facilidad de obtener resultados de datos considerados complicados y no muy precisos, se utilizan para extraer patrones y hallar acertadamente clasificaciones de datos que son muy difíciles con otras técnicas (Maynard, 2020). Según Kohonen, (1989), "Las RNA son grupos de elementos de cálculo sencillas, adaptativas, que se interconectan de forma masiva paralelamente, organizadas jerárquicamente que facilita la interacción en el sistema de la misma forma que lo hace el sistema biológico.

La facilidad del aprendizaje adaptativo que muestran, el proceso de organización, la tolerancia a fallos que presentan, el manejo de operaciones en tiempo real y la capacidad de inserción dentro de las tecnologías existentes, han logrado que su utilización sea de amplio uso en áreas como la clasificación biológica (Chen et al., 2019; Li et al., 2019), en procesos financieros (Lahmiri & Bekiros, 2019; Sushant K, 2020), en procesos industriales (Elhoone et al., 2020; Hesser & Markert, 2019), en aspectos y proceso del medio ambiente (Loy-Benitez et al., 2020; Tuttle et al., 2019), en la detección de fallos en sistemas de

control (Wang et al., 2020; Wang et al., 2020), en salud (Mehanović et al., 2020; Morrow et al., 2020) y optimización de procesos (Schweidtmann & Mitsos, 2019; Zhou et al., 2020).

- Red neuronal artificial definición

Una red neuronal es un grafo dirigido que tiene las siguientes propiedades:

- Cada neurona i tiene asociada una variable de estado X_i .
 - Cada conexión (i, j) entre las neuronas i y j tiene asociado un peso $w_{ij} \in \mathbb{R}$.
 - Cada neurona i tiene asociado un umbral $\theta_i \in \mathbb{R}$.
 - Cada neurona i tiene definida una función $f_i(x_1, \dots, x_n, w_{i1}, \dots, w_{in}, \theta_i)$ dependiente de los pesos de las conexiones, del umbral y del estado de cada neurona j que estén conectadas con la neurona i . El valor de esta función proporciona el nuevo estado de la neurona.
- Neurona artificial y su modelo estándar

Según los principios enunciados por los autores (McClelland & Rumelhart, 1986; Rumelhart & MacClelland, 1986) la i -ésima neurona artificial funciona de la siguiente manera:

- a. El conjunto de entradas dado X_j y los pesos sinápticos W_{ij} , $\rightarrow j = 1, \dots, n$
- b. Regla de propagación h_i que se define a partir del conjunto de entradas y con los respectivos pesos sinápticos. Es decir: $h_i(x_1, \dots, x_n, w_{i1}, \dots, w_{in})$.

La regla de propagación de mayor utilización resulta de combinar linealmente las entradas y los pesos sinápticos, de donde se obtiene:

$$h_i(x_1, \dots, x_n, w_{i_1}, \dots, w_{i_n}) = \sum_{j=1}^n w_{ij} x_j \quad (1)$$

Es común adicionar a los pesos de la neurona un parámetro denominado θ_i , el cual se reconoce como umbral, y que se acostumbra a restar al potencial pos-sináptico. Es decir:

$$h_i(x_1, \dots, x_n, w_{i_1}, \dots, w_{i_n}) = \sum_{j=1}^n w_{ij} x_j - \theta_i \quad (2)$$

Si se hace que los índices i y j comiencen en 0, y denotamos por $w_{i_0} = \theta_i$ y $x_0 = -1$, la regla de propagación se puede expresar:

$$h_i(x_1, \dots, x_n, w_{i_1}, \dots, w_{i_n}) = \sum_{j=0}^n w_{ij} x_j = \sum_{j=1}^n w_{ij} x_j - \theta_i \quad (3)$$

- c. La función de activación, la cual representa simultáneamente la salida de la neurona y su estado de activación. Si denotamos por y_i dicha función de activación, tenemos:

$$y_i = f_i(h_i) = f_i\left(\sum_{j=0}^n w_{ij} x_j\right) \quad (4)$$

- d. La neurona continua sigmoidea, la cual permite obtener una salida continua, la más usada es:

$$y_i = \frac{1}{1+e^{-\sum_{j=1}^n w_{ij}x_j - \theta_i}}, \text{ con } y_i \in [0, 1] \quad (5)$$

- Arquitecturas de redes neuronales

Se le llama arquitectura a la estructura de una red neuronal. Las redes neuronales están constituidas por nodos que se conectan por medio de sinapsis, el comportamiento de la red está dado por la estructura de conexiones sinápticas. Estas conexiones propagan o comunican la información en un solo sentido, (de la neurona pre sináptica a la pos-sináptica), las neuronas por lo general se agrupan en unidades denominadas capas.

El conjunto total de capas conforma la red neuronal. Generalmente, las redes poseen una capa de entrada, una de salida y unas capas ocultas. Las de entrada reciben las señales, las de salida dan a conocer la respuesta de la red y las capas ocultas proporcionan grados de libertad a la red con los cuales representa las características de la realidad que se quiere modelar.

Las redes presentan diferentes arquitecturas, entre ellas las monocapa y las multicapa. las redes acordes al flujo de datos se clasifican en feedforward o unidireccionales y feedback o realimentadas. En las redes realimentadas los datos pueden circular entre las distintas capas, sin importar el sentido.

- Asociador Lineal

El rendimiento de una red neuronal a partir de una regla de actualización de pesos que la logre optimizar, se puede hallar a través del error cuadrático medio de las salidas con respecto a las deseadas. El asociador lineal calcula el potencial pos-sináptico por medio de la convencional suma ponderada, cantidad a la que aplica posteriormente una función de activación de tipo identidad. El asociador lineal debe aprender a asociar N pares entrada-salida ajustando los pesos para luego calcular el porcentaje correcto de las salidas con el valor real. Dónde:

$$\frac{1}{N} \sum_{r=1}^N \|y^r - w x^r\| = \frac{1}{N} \sum_{r=1}^N \sum_{i=1}^m \sum_{j=1}^n (y_i^r - W x_j^r)^2 \quad (6)$$

- Perceptrón Multicapa de Retropropagación (Multi-Layer Perceptron BackPropagation)

Este tipo de redes neuronales se utiliza gracias a que la implementación de múltiples capas en su diseño facilita la identificación de regiones convexas y discriminar regiones de forma aleatoria. La red MLP suele utilizar algoritmos de Retropropagación de error para facilitar el procesamiento de los datos. Las redes MLP utilizan funciones de tipo sigmoideo (Werboz, 1974).

2.2.4 Métricas de evaluación del desempeño

Existen varias métricas de rendimiento que se utilizan para determinar la calidad de un clasificador. La más común es una matriz de confusión. Una matriz de confusión contiene el rendimiento de clasificación del algoritmo, es decir, los datos reales versus la clase predicha alcanzada por el clasificador en un conjunto de muestras (Han et al., 2012) (

Tabla 2 - 1).

Tabla 2 - 1. Matriz de confusión.

		Predicción			
		Negativo	Positivo		
Real	Negativo	a	B		
	Positivo	c	D	Exactitud	$TP + TN / TP + TN + FP + FN$
		Sensibilidad	Especificidad		
		$TP / TP + FN$	$TN / TN + FP$		

Total positivos (TP): Número de casos que la prueba declara positivos y que son verdaderamente positivos.

Falsos positivos (FP): Número de casos que la prueba declara positivos y que en realidad son negativos.

Total negativos (TN): Número de casos que la prueba declara negativos y que son realmente negativos.

Falsos negativos (FN): Número de casos que la prueba declara negativos y que en realidad son positivos.

Métricas derivadas de la matriz de confusión para evaluar el algoritmo.

$$\text{Sensibilidad} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{Especificidad} = \frac{TN}{TN + FP} \quad (8)$$

$$\text{Error de Clasificación} = \frac{FP + FN}{TP + TN + FP + FN} \quad (9)$$

$$\text{Exactitud} = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

2.2.5 Curva ROC

Los gráficos ROC son gráficos bidimensionales en los que la tasa Tp se traza en el eje Y y la tasa Fp se traza en el Eje X. Un gráfico ROC representa compensaciones relativas entre beneficios (verdaderos positivos) y costos (falsos positivos). Cada clasificador produce un par de puntos (tasa Fp , tasa Tp) correspondiente a un solo punto en el espacio ROC (Fawcett, 2006).

2.3. Aplicación de las técnicas de Inteligencia Artificial en el campo de la salud y de la Seguridad y Salud en el Trabajo

El problema de la evaluación y la clasificación del riesgo psicosocial ha sido estudiado a través de las técnicas de inteligencia artificial, algunos autores han desarrollado aproximaciones en trabajos donde a parte de la clasificación del nivel de riesgo psicosocial, se asocia los niveles de estrés, estrés mental, salud mental, satisfacción laboral, rendimiento laboral y síntomas generales de salud. La siguiente recopilación bibliográfica se centra en los trabajos que desde el aprendizaje automático y la minería de datos disponible en las bases de datos realizan una aproximación a la identificación de los factores y del grado de riesgo psicosocial.

Desde la Seguridad y la Salud en el Trabajo la evaluación de riesgos ha ido implementando las técnicas de aprendizaje de máquina para ayudar a los prevencionistas y a los encargados de la seguridad en los ambientes laborales a mejorar la identificación de los riesgos, varios procesos de manufactura y de la industria en general han contribuido a esto haciendo uso de las redes neuronales artificiales, como es el caso de la industria automotriz (Hegde & Rokseth, 2020), pero muy pocos se han centrado en la identificación de los riesgos psicosociales en los procesos de servicios y manufactura.

Posada-Quintero et al., (2020) realizaron un análisis exhaustivo en 57 docentes de colegios en Colombia para observar los factores de riesgo y síntomas asociados a burnout (síndrome del quemado), para ellos utilizaron técnicas de máquinas de soporte vectorial y árboles de decisión.

La aplicación de los modelos permitió identificar que la insatisfacción con el ingreso del trabajo era el factor de riesgo más relevante, seguido por el trabajo de horas extras y la severidad percibida de las sanciones en el desempeño inferior. Los síntomas más relevantes del agotamiento fueron, fatiga al final del día y frecuentes y recurrentes dolores de cabeza. Recomiendan utilizar la herramienta o el desarrollo de éstas en la identificación de problemas asociados a factores psicosociales.

González Fuentes et al., (2020) señalan lo siguiente, basados en la sexta encuesta europea de condiciones del trabajo: El uso de técnicas de aprendizaje automático para predecir resultados de salud y seguridad en el trabajo en diferentes campos no es nuevo, ya sea que se centre en accidentes laborales, riesgo de incendio, MSD o trastornos visuales. Algunos de estos trabajos se centran en sectores específicos, como la minería o la industria de la salud. Sin embargo, hasta la fecha, ninguna investigación ha estudiado cómo la combinación de factores como la edad, el género, el bienestar, la vida doméstica y social, así como los factores psicosociales, pueden influir en la propensión a las bajas por enfermedad entre los trabajadores del sector energético. Ellos aplicaron un modelo híbrido de regresión adaptativa multivariada por tramos (MARS) y algoritmos genéticos para clasificar y predecir el número de días de ausentismo laboral debido a las variables mencionada anteriormente.

Uronen et al., (2020) propusieron la detección automatizada del riesgo psicosocial en enfermeras declarando que, los factores de riesgo psicosocial influyen en la jubilación anticipada y la ausencia del trabajo. Realizaron el análisis a través del análisis de minería de texto en registros nacionales de documentación llenada por enfermeras ocupacionales. Su análisis se basó teniendo en cuenta los 13 factores psicosociales que evalúa el estándar Prima – EF de riesgos psicosociales (PAS1010) (Jain et al., 2017) y utilizaron un modelo de predicción basado en redes neuronales adaptado denominado Word2vec (Mikolov et al., 2013) para identificar las palabras asociadas a posibles problemas psicosociales.

Márquez Gómez, (2020) aplicó los modelos de regresión logística y árboles de decisión para determinar el nivel de discomfort en hombros, espalda, brazos, manos y cuello asociados a factores psicosociales, los resultados estuvieron entre un 83.3% y 90.2% de exactitud de la clasificación.

Al-Shargie, (2019) menciona como el estrés mental ha sido identificado como uno de los principales factores que contribuye a producir ataques cardíacos, depresión y accidentes cerebro vasculares (ACV). Aplicó máquinas de soporte vectorial junto a código de salida de corrección de errores (ecoc) utilizando señales de electroencefalografía para diferenciar diferentes niveles de estrés, según su modelo de aritmética mental (MA), obteniendo resultados de clasificación alrededor de 94,79%.

Mosquera et al., (2019) utilizaron un modelo híbrido de k vecinos más próximos junto a máquinas de soporte vectorial (K-nn-SVM) para modelar y clasificar los riesgos psicosociales comparándolo con los modelos individualmente de k-vecinos más próximos y de máquinas de soporte vectorial lineal, obteniendo un resultado de clasificación del 86,66%.

Jebelli et al., (2019) utilizaron dos redes neuronales de aprendizaje profundo basadas en señales EEG para reconocer el estrés en los trabajadores de la construcción, con un 86.82% de exactitud en comparación con los métodos anteriores de reconocimiento de estrés.

Jung et al., (2019) utilizaron técnicas algorítmicas como random forest y un extreme gradient boosting entre otras técnicas para la predicción de alto riesgo de suicidio en adolescentes coreanos, identificaron la violencia y el estrés como

factores asociados. Utilizaron 26 variables como predictores, para predecir el suicidio de alto riesgo. La exactitud fue mejor en el aumento de gradiente extremo (79.0%).

Un análisis realizado por Mosquera, Castrillón, et al., (2018b) comparando tres modelos de aprendizaje supervisado para la clasificación: redes neuronales artificiales, arboles de decisión y Naïve Bayes, utilizando reducción de la dimensionalidad de los datos, aplicaron algoritmos genéticos, algoritmo de la cantidad esperada de la información, algoritmo GainRatioAttributeEval, coeficiente de relación de Pearson y análisis principal de componentes. Se utilizó una base de datos con 5340 registros de pacientes, que correspondían a evaluaciones psicosociales de docentes de colegios públicos de Colombia. El mejor desempeño de clasificación lo obtuvo redes neuronales artificiales con una efectividad del 93%.

Un modelo para evaluar la efectividad en la clasificación de riesgo psicosocial fue desarrollado por Mosquera, Gómez, et al., (2018) implementando un modelo de Hill Climbing - Support Vector Machine (HC-SVM) logrando optimizar los valores de entrada al modelo para optimizar los parámetros de clasificación, se incluyeron tres nuevas variables en el estudio como, presión arterial, electromiografía y actividad eléctrica de la piel en una muestra de 480 individuos voluntarios. El porcentaje de clasificación que se logró fue del 95%.

Mosquera, Castrillón, et al., (2018c) realizaron un estudio comparativo del modelo de aprendizaje de máquina para la predicción utilizando Máquinas de Soporte Vectorial (SVM) y Naïve Bayes, en dos etapas, primero con todas las variables psicosociales y segundo, reduciendo la dimensionalidad de los datos utilizando algoritmos genéticos. Se seleccionaron las cuarenta mejores variables con mejor eficiencia en la exactitud de la predicción. La base de datos utilizada se componía

de 3000 registros epidemiológicos, que correspondían a docentes de colegios públicos del área metropolitana de una ciudad colombiana.

Tiffin & Paton, (2018) proponía: el aprendizaje automático no debe considerarse una solución mágica para los problemas predictivos en la atención de la salud mental, sino como una herramienta en el cuadro cuantitativo de enfoques que pueden ser útiles tanto en la investigación como, eventualmente, en la práctica.

Al igual que una clase novedosa de drogas psicotrópicas, el aprendizaje automático indudablemente finalmente encontrará su lugar correcto dentro de los servicios de salud mental, dejando claro que, el aprendizaje automático tiene el potencial de abordar problemas importantes de diagnóstico y toma de decisiones clínicas en salud mental. Desde ésta perspectiva varios autores han trabajado en el desarrollo de aplicaciones que permitan contribuir a la identificación de los riesgos psicosociales.

Yadav & Hashmi, (2018) aplicaron algoritmos como support vector machine (SVM), neural network (NN), decision tree (DT) and random forest (RF) para detectar el estrés ocupacional en empleados del sector seguros y financiero, encontrando resultados prometedores a través del algoritmo random forest obtuvo una exactitud del 60%, una sensibilidad del 80% y una especificidad del 60%. Hallaron que para éste estudio la evaluación del estrés ocupacional está críticamente conectada al desempeño laboral en la organización.

Lal et al., (2018) aplicaron un detector de estrés propuesto que diferencia a una persona estresada de una normal, al obtener sus señales fisiológicas, utilizando sensores como electrocardiograma, sensores de frecuencia cardíaca y señales psicológicas a través de los puntajes de la evaluación de socio-estrés. El resultado ayuda a detectar el factor que causó el aumento del nivel de estrés. Las señales se procesan para clasificar las características que indican el nivel de estrés en

personas que trabajan. Utilizaron máquinas de soporte vectorial para clasificar estos conjuntos de características adquiridos logrando buenos resultados.

Anand & Kumar, (2018) utilizaron señales de electroencefalograma para asociarlo a estrés mental de 41 pacientes voluntarios e identificarlo, comparándolo con un cuestionario médico para observar los aciertos, utilizaron un perceptrón multicapa junto a máquinas de soporte vectorial como kernel para realizar la clasificación de las señales.

García-Herrero et al., (2017) utilizaron una red bayesiana en profesionales de la salud con la cual hallaron que las demandas emocionales tienen una mayor influencia en el aumento de la probabilidad de estrés debido a la carga de trabajo que las demandas familiares. Los resultados muestran que el apoyo de los colegas, en general, tiene menos efecto en la reducción del estrés que el apoyo social de los superiores. Además, el análisis de sensibilidad muestra que, en situaciones de alta demanda y bajo control, el reconocimiento claramente impacta el estrés, reduciéndolo drásticamente.

Darvishi et al., (2017) utilizaron un modelo de predicción del dolor lumbar (LBP) considerando las interacciones de factores de riesgo ocupacionales, personales y psicosociales en población de trabajadores en unidades industriales utilizando un enfoque de redes neuronales artificiales. El promedio de exactitud de clasificación de las redes neuronales desarrolladas para los datos de la fase de prueba y entrenamiento fue de aproximadamente 88% y 96%, respectivamente. Además, la media de exactitud de clasificación de los datos de entrenamiento y evaluación fue del 92%, lo que indica resultados mucho mejores en comparación con otros métodos.

Una investigación realizada por Li et al., (2017) utilizó modelos optimizados de redes neuronales para detectar cambios en las redes organizacionales, utilizaron

una técnica basada en redes neuronales (BPNN) con algoritmos genéticos (GA) y optimización de enjambre de partículas (PSO). Los resultados del estudio sugirieron que el método propuesto podría identificar los factores psicosociales presentes en las organizaciones laborales y puede cambiar el clima organizacional de las empresas con valores de curva ROC entre 84% y 87%.

Subhani et al., (2017) han utilizado diferentes técnicas inteligentes (regresión logística, máquina de vectores de soporte y clasificadores ingenuos de Bayes) para la detección de estrés mental en múltiples niveles, y para dos niveles, los resultados, respectivamente, mostraron que la propuesta produjo una exactitud del 83.4% para la identificación de múltiples niveles y una exactitud del 94.6% para la identificación de estrés en dos niveles. Impulsaron y recomendaron el desarrollo de una herramienta de diagnóstico por computadora para la detección de estrés.

Mosquera et al., (2016) aplicaron minería de datos en diagnósticos de evaluaciones psicosociales para identificar el nivel de riesgo psicosocial en docentes de colegios públicos en Colombia. Tomaron una muestra de 1100 datos, correspondientes a evaluaciones psicosociales. Compararon los resultados obtenidos al aplicar algoritmos de árboles de decisión J48 y Naive Bayes. Obtuvieron una efectividad del 91% en la clasificación.

González Herranz, (2016), desarrolló un sistema para facilitar el diagnóstico y determinar el tipo de terapia funcional que se pueda emplear para enfermedades definidas como neurodegenerativas, el cual utilizó clasificadores inteligentes: redes neuronales (MLP y LVQ) y máquinas de soporte vectorial (SVM), las cuales identifican los tipos de temblor que se pueden asociar al Parkinson, para poder determinar con antelación el tipo de terapia y tratamiento adecuado que se le dará al paciente. Decidió realizar la tesis doctoral en éste ámbito pues, se identificó

que actualmente los profesionales de la salud, utilizan escalas cualitativas para poder identificar la patología y el grado de afectación de la misma.

La aplicación de redes bayesianas para el cálculo de probabilidad de sufrir estrés asociado a problemas de salud en un estudio que se llevó a cabo en España teniendo como base la V encuesta europea de condiciones de trabajo, destacan que la probabilidad más alta de sufrir niveles de estrés elevados la tienen cuando se introduce la variable del insomnio en el modelo. Es decir, si no se es capaz de conciliar el sueño regularmente, esas probabilidades de sufrir niveles de estrés elevados ascienden a un 15,58% (Cardenas, 2015), el valor de exactitud de la implementación del modelo fue del 62% para identificar sintomatología de estrés laboral.

Galatzer-Levy et al., (2014) evalúan la capacidad de predicción de los enfoques de aprendizaje de máquina (ML) utilizando varios modelos de SVM para lograr identificar e integrar características únicas predictivas y determinar su exactitud en la predicción de la remisión de trastorno de estrés postraumático a partir de información recopilada en 10 días de ocurrido un evento traumático. Los datos sobre las características de los eventos, las observaciones del servicio de urgencias y los primeros síntomas se recogieron en 957 sobrevivientes de trauma, seguidos durante quince meses.

La capacidad de identificar e integrar los indicadores de riesgo, hacen de este un enfoque prometedor para el desarrollo de algoritmos que infieren probabilísticamente el riesgo de la psicopatología de estrés postraumático crónico basado en fuentes complejas de información biológica, psicológica y social. Los resultados de exactitud que lograron se encontraban entre el 81 y 82%.

Ye et al., (2014) también muestran que las técnicas inteligentes son útiles para el reconocimiento de patrones, en las que el comportamiento mental, emocional y

fisiológico del ser humano está implicado. Por ejemplo, señalan que, según los resultados de los experimentos, seis sentimientos o emociones básicas (disgusto, ira, miedo, tristeza, alegría y sorpresa), alto estrés mental y movimientos irregulares pueden ser reconocidos por los algoritmos con altas precisiones. Así, los tres tipos de reconocimiento de patrones se pueden utilizar para la detección de acoso escolar (bullying) con eficacia.

Wenhui Liao et al., (2005) han utilizado las Redes Bayesianas Dinámicas (DBN) para intentar el reconocimiento de los patrones asociados con el estrés laboral. Esto se realiza a través del reconocimiento variable, que incluye expresiones faciales, movimientos oculares y movimientos de la cabeza extraídos de videos previamente grabados de sensores visuales, condiciones fisiológicas tomadas de un mouse emocional, datos de comportamiento de las actividades del usuario en las que interactúan con la computadora y medición de rendimiento. El estudio se realizó con cinco sujetos con objeto de estudio. El alto nivel de estrés se mostró a través de los síntomas, incluidos latidos cardíacos rápidos, respiración rápida, aumento de la sudoración, piel fría, manos y pies fríos, náuseas y músculos tensos, entre otros. Los mejores resultados de correlación se obtuvieron con el tercer sujeto que mostró una correlación combinada entre estrés y las variables de 0.92.

Tzeng et al., (2004) realizaron una búsqueda en la base de datos Medline, encontrando que las máquinas de soporte vectorial no habían sido previamente utilizadas en el área de gestión de recursos humanos. Se plantaron la hipótesis de que este enfoque podría ser útil para establecer un mecanismo de alerta temprana para la administración de recursos de enfermería. Decidieron realizar este estudio basado en estudios anteriores que indicaban que la satisfacción laboral, el compromiso de la organización, la motivación para realizar bien el trabajo, las dimensiones cooperativas dentro del trabajo, la complejidad del trabajo, la ayuda o soporte recibido de parte de los jefes y el tiempo suficiente para el cuidado personal, influían en la decisión de continuar en su trabajo.

El estudio se llevó a cabo en Taiwán en 3 hospitales con enfermeras, donde aplicaron un cuestionario para obtener los datos, los que tenían como objetivo ser utilizados para entrenar una máquina de aprendizaje (SVM) para predecir la intencionalidad de las enfermeras de renunciar al trabajo.

Los predictores incluyeron datos de las características demográficas, la motivación en el trabajo, los niveles de satisfacción en el trabajo y los niveles de estrés, los cuales se identificaron como estresores ocupacionales. Se identificaron 389 casos que respondieron correctamente los cuestionarios distribuidos en los 3 hospitales de mayor influencia en la ciudad, de los cuales se obtuvo la información para entrenar la máquina de soporte vectorial. Para validar los datos, realizaron validación cruzada, dividiendo de manera aleatoria en 4 grupos de aproximadamente el mismo tamaño y realizaron 4 corridas de entrenamiento. Después de éste, el porcentaje medio de la clasificación errónea de los datos de entrenamiento fue de 0,86, mientras que en los datos de prueba fue de 10,8, lo que resultó en predicciones con el 89,2% de exactitud. Al realizar varias pruebas de validación del algoritmo, los autores afirman que esta máquina de soporte vectorial puede ayudar a predecir la intencionalidad de dejar el trabajo por parte de las enfermeras, sin necesidad de preguntarles a ellas.

Hallner & Hasenbring, (2004) utilizaron una red neuronal perceptron multicapa de tres capas para predecir cronicidad en pacientes con dolor lumbar a través de 28 variables psicosociales ya declaradas, que permiten identificar dolores tempranos asociados a la ocurrencia de factores psicosociales.

Kataoka et al., (1998) han desarrollado un sistema que mide la temperatura corporal de un individuo en el lugar de trabajo, para asociarlo con el estrés laboral, y han aplicado una ecuación de regresión lineal, para predecir el nivel de estrés

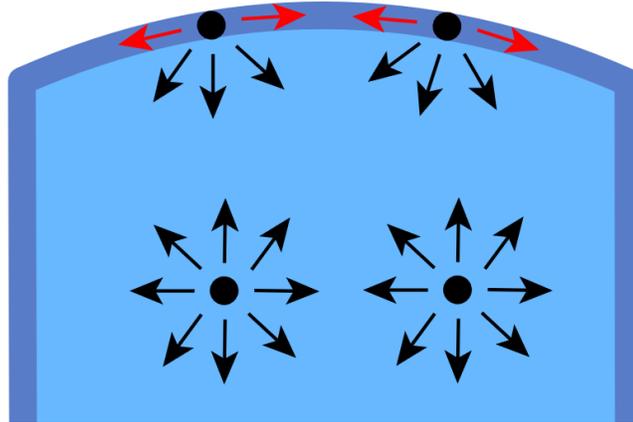
del trabajador con respecto a las tareas que deben realizarse. Proponen mejorar las tecnologías y asociarlo a métodos cuantitativos.

2.4. Teoría de la Tensión Superficial

En las ciencias físicas, por definición la tensión superficial de un líquido es la cantidad de energía que se necesita para aumentar la superficie por unidad de área. Los líquidos presentan una resistencia o una fuerza contraria a la ejercida sobre ellos en su superficie, la cual se denomina fuerza intermolecular, es la que permite a ciertos insectos o arácnidos desplazarse sobre el agua.

De la forma más simple, la tensión superficial se da porqué las fuerzas intermoleculares del líquido al interior y en la superficie son diferentes. Al interior cada molécula está atraída por fuerzas que hace que se mantengan unidas, pero con una cantidad de energía baja, en la superficie del líquido la fuerza que se ejerce hacia el interior del líquido trata de romper dicha fuerza molecular interna.

Otra forma de explicarlo sería, cada molécula tiene contacto con su vecina en estados similares de energía, Las moléculas vecinas entre sí tienen toda la mayor cantidad posible de moléculas que pudieran tener, manteniendo el sistema, pero las partículas que llegan desde el exterior tienen menos partículas vecinas con un estado energético más alto, las moléculas al interior tratan de estabilizar el sistema para evitar el rompimiento de la capa superficial, esto, buscando disminuir el estado energético del exterior (**Figura 2- 3**).

Figura 2- 3. Fuerzas entre dos moléculas de un líquido.

Fuente: Tomada de

<https://commons.wikimedia.org/wiki/File:Wassermolek%C3%BCleInTr%C3%B6pfchen.svg>

Por lo tanto, las fuerzas que se generan a nivel molecular en un líquido que permite que la superficie no se rompa, se denominan cohesivas, dando como tal el fenómeno de tensión superficial que busca estabilizar el sistema de fuerzas (Ecuación 8) (Adamson & Gast, 1967; Berry, 1971, p. 1946; Brown, 1947; Fowkes, 1961; Law & Zhao, 2016; Navascués, 1979).

$$\gamma = \frac{F}{2L} \quad (11)$$

Dónde,

γ es la tensión superficial de un líquido por unidad de longitud.

F es la fuerza requerida para sujetar el lado móvil, que es proporcional a la longitud L del lado inmóvil.

L es la longitud del lado móvil.

$\frac{1}{2}$ debido a que la película del líquido tiene dos lados (dos superficies), cada una de las cuales contribuye por igual a la fuerza (trata de igualar las fuerzas en el sistema).

2.5. Factores de riesgo psicosocial

Las condiciones sociales que afectan la salud del trabajador en su parte psíquica, física y emocional, se empezaron a estudiar desde el año 1948, cuando los trabajadores de un restaurante manifestaron sufrir de síntomas de tensión graves que les ocasionaba el atender las demandas de altos volúmenes de clientes que no se pudieron controlar y tenían relación directa con las actividades laborales propias de la función, ligadas a la falta de una adecuada organización del trabajo dentro de las funciones que debía desempeñar cada trabajador (Karasek, 1979). Desde ese entonces, muchos modelos se han desarrollado para estudiar la relación factores psicosociales salud del trabajador.

2.5.1. Modelo de demanda, control, apoyo social

El modelo de demandas-control apoyo social, inicialmente fue desarrollado por (Karasek, 1979), el cual estudiaba los efectos del rediseño del puesto de trabajo, sumado a las variables de estudio de autonomía laboral, satisfacción, motivación y compromiso, teniendo en cuenta los efectos de éstas sobre la salud del trabajador.

Johnson & Hall, (1988) ampliaron la teoría aportando que las variables cansancio, depresión, ansiedad, percepción de estrés, percepción de salud, accidentes laborales y bajas por enfermedad estaban ligadas a la afectación de la salud mental del trabajador en el ambiente laboral y como el apoyo social de los compañeros de trabajo ayudaba a disipar la presión psicológica (**Figura 2- 4**).

Según el modelo, plantea que, el poder llegar a un equilibrio entre los tres factores, depende específicamente de la organización del trabajo, por lo cual, se centra únicamente en las condiciones laborales que pueden afectar al individuo y de forma específica en las características psicosociales del ambiente laboral. Este modelo es el que mayormente se ha utilizado para explicar con evidencia científica la relación estrés-efectos en la salud.

2.5.2. Modelo de desequilibrio esfuerzo-recompensa

El modelo desarrollado por Siegrist, (1996) explica la manera en que el estrés laboral afecta la salud del trabajador debido a la falta de control que se tiene sobre su futuro laboral, las recompensas laborales a largo plazo, la amenaza de perder su trabajo por despido, la baja estabilidad laboral debido a malos contratos, la inconsistencia laboral y el poco y mal desarrollo profesional en ámbitos laborales precarios.

El modelo trata de explicar como las consecuencias diarias laborales del estrés y de las características psicosociales del entorno laboral produce alteraciones de orden fisiológico en órganos como corazón, intestinos, músculos, vasos sanguíneos, cerebro y como ante situaciones que se pueden denominar de frustración ante no recibir recompensas en las actividades laborales, producen una tensión sostenida en el individuo, desencadenando enfermedades laborales.

Se considera en éste modelo que los factores psicosociales que conducen a estrés constante en las actividades laborales, desencadenan también la activación del sistema nervioso del individuo.

Figura 2- 4. Modelo teórico de exigencias psicológicas- demandas-control apoyo social



Fuente: Adaptado de: (Johnson & Hall, 1988; Karasek, 1979).

2.5.3. Modelo dinámico de los factores de riesgo psicosocial

El modelo propuesto por Villalobos, (2005) se basa en los modelos de Karasek y Theorell y Siegrist, el cual postula que los ambientes de trabajo son un sistema en el cual las relaciones y los componentes que lo conforman son cambiantes y estos generan estrés que afectan la salud del trabajador.

2.5.4. Condiciones Intralaborales de los riesgos psicosociales

Se consideran a las condiciones intralaborales (**Tabla 2 - 2**) a aquellas características y variables sistémicas presentes en la organización del trabajo que afectan la salud y el bienestar del trabajador. Las condiciones extralaborales, son aquellas que comprenden el entorno de vivienda, familiar, social y económico del trabajador y que pueden influir en su salud y las condiciones individuales, son aquellos factores como el estado civil, la ciudad o zona de residencia, el nivel escolar, el sexo, la edad, el número de personas que dependen del trabajador, el estrato socio-económico en el que vive el trabajador, el tipo de trabajo que desempeña.

Por ser de importancia para el tipo de formación doctoral que se recibe y por la profesión del investigador, ésta tesis doctoral se centró en las condiciones intralaborales, que, desde la organización del trabajo, el diseño de las operaciones, el diseño del trabajo, el diseño de las condiciones laborales, el diseño macroergonómico de las condiciones laborales, el diseño de las condiciones de Seguridad y Salud en el Trabajo y desde el desarrollo y la aplicación de la inteligencia artificial, son temas que se pueden abordar.

La batería de riesgo psicosocial como instrumento de desarrollo por profesionales de la salud, de la psicología, de la estadística con conocimiento en el diseño de instrumentos

psicotécnicos y basado en la revisión de literatura de 110 artículos científicos sobre métodos e instrumentos de evaluación de factores psicosociales.

“El modelo dinámico de los factores psicosociales identifica cuatro dominios que agrupan un conjunto de dimensiones que explican las condiciones intralaborales. Las dimensiones que conforman los dominios actúan como posibles fuentes de riesgo (Villalobos, 2005) y es a través de ellas que se realiza la identificación y valoración de los factores de riesgo psicosocial” (Villalobos et al., 2010, p. 20).

- Demandas del trabajo: Hace referencia a las exigencias que todo individuo tiene por el trabajo, pueden ser de tipo cuantitativas, cognitivas, mentales, emocionales, de responsabilidad, debido a la jornada de trabajo o por el ambiente físico del lugar de trabajo.
- Control sobre el trabajo: Hace referencia a que un individuo pueda tomar decisiones en su aspecto laboral, aplicar el conocimiento y habilidades en tareas de manera autónoma, participar en decisiones y tener claro su rol en la compañía, lo que posibilita que el trabajador pueda tener influencia en su trabajo.
- Liderazgo y relaciones sociales en el trabajo: es la relación que exista entre los jefes y los trabajadores, y como ésta relación puede influir en la forma de desempeñarse y en el clima organizacional de un área de trabajo.
- Recompensa: Son aquellas retribuciones que se obtienen por el esfuerzo de parte del trabajador en su trabajo para ayudar en el desarrollo de la compañía, pueden ser de tipo económico, social, educativo, de bienestar, de promoción en el trabajo, de planes de salud, entre otros. (Villalobos et al., 2010, pp. 23–25)

2.5.5. Conjunto de datos utilizado para este trabajo

Se consideraron 114 variables, con valores cuantitativos y cualitativos y 5443 registros de muestra correspondientes a la aplicación de la batería de riesgos psicosocial intralaboral (Ministerio de la Protección Social, 2010) en docentes de colegios públicos del área metropolitana de una ciudad colombiana (Cali, Palmira, Yumbo, Jamundí y Candelaria), la aplicación de la evaluación se hizo de manera directa a cada uno de las personas (**Tabla 2 - 2**), la que contenía 114 variables predictoras cuantitativas y 1 variable de clase (riesgo) (Mosquera, 2019; Mosquera, 2018).

Se evaluaron también, las condiciones musculo esqueléticas y fisiológicas en las historias clínicas de pacientes que declaraban tener alguna afección y se realizó una muestra a algunos pacientes que decidieron participar voluntariamente y se realizó valoración por parte médica. Las afecciones y las variables aquí denotadas aparecieron al momento de la entrevista con los docentes, muchos de ellos declaraban haber tenido algún tipo de afección musculoesquelética o de problemas relacionados con presión arterial o malestar en el pecho, dolor en alguna parte del cuerpo, dolores de cabeza en algún momento, estos síntomas son reconocidos como afecciones somáticas crónicas en estudios anteriores (Dimsdale, 2019; Sauter & Murphy, 1984), estos síntomas declarados por parte de los docentes, llevaron a tomar en cuenta éstas variables para el análisis en esta investigación como variables que en algún momento aparecían en los individuos debido a alguna situación presentada en el ambiente laboral (**Tabla 2 - 3 – Tabla 2 - 4**).

Tabla 2 - 2. Variables de los dominios de factores de riesgo psicosocial intralaboral. Tomado de: Batería para la evaluación de factores de riesgo psicosocial (Villalobos et al., 2010).

No.	Dominios	Variable
S	Sociodemográficas	Sexo (S1), edad(S2), casado(S3), unión libre(S4), soltero(S5), posgrado(S6), estrato(S7), tipo de vivienda(S8), número de personas que dependen económicamente(S9), antigüedad en el cargo(S10).
x1	Demandas del Trabajo	El ruido en el lugar donde trabajo es molesto
x2		En el lugar donde trabajo hace mucho frío
x3		En el lugar donde trabajo hace mucho calor
x4		El aire en el lugar donde trabajo es fresco y agradable
x5		La luz del sitio donde trabajo es agradable
x6		El espacio donde trabajo es cómodo
x7		En mi trabajo me preocupa estar expuesto a sustancias químicas que afecten mi salud
x8		Mi trabajo me exige hacer mucho esfuerzo físico
x9		Los equipos o herramientas con los que trabajo son cómodos
x10		En mi trabajo me preocupa estar expuesto a microbios, animales o plantas que afecten mi salud
x11		Me preocupa accidentarme en mi trabajo
x12		El lugar donde trabajo es limpio y ordenado
x13		Por la cantidad de trabajo que tengo debo quedarme tiempo adicional
x14		Me alcanza el tiempo de trabajo para tener al día mis deberes
x15		Por la cantidad de trabajo que tengo debo trabajar sin parar
x16		Mi trabajo me exige hacer mucho esfuerzo mental
x17		Mi trabajo me exige estar muy concentrado

No.	Dominios	Variable
x18	Demandas del Trabajo	Mi trabajo me exige memorizar mucha información
x19		En mi trabajo tengo que tomar decisiones difíciles muy rápido
x20		Mi trabajo me exige atender a muchos asuntos al mismo tiempo
x21		Mi trabajo requiere que me fije en pequeños detalles
x22		En mi trabajo respondo por cosas de mucho valor
x23		En mi trabajo respondo por dinero de la empresa
x24		Como parte de mis funciones debo responder por la seguridad de otros
x25		Respondo ante mi jefe por los resultados de toda mi área de trabajo
x26		Mi trabajo me exige cuidar la salud de otras personas
x27		En el trabajo me dan órdenes contradictorias
x28		En mi trabajo me piden hacer cosas innecesarias
x29		En mi trabajo se presentan situaciones en las que debo pasar por alto normas o procedimientos
x30		En mi trabajo tengo que hacer cosas que se podrían hacer de una forma más práctica
x31		Trabajo en horario de noche
x32		En mi trabajo es posible tomar pausas para descansar
x33		Mi trabajo me exige laborar en días de descanso, festivos o fines de semana
x34		En mi trabajo puedo tomar fines de semana o días de descanso al mes
x35		Cuando estoy en casa sigo pensando en el trabajo
x36		Discuto con mi familia o amigos por causa de mi trabajo
x37		Debo atender asuntos de trabajo cuando estoy en casa
x38		Por mi trabajo el tiempo que paso con mi familia y amigos es muy poco
x39		Puedo tomar pausas cuando las necesito
x40		Puedo parar un momento mi trabajo para atender algún asunto personal
x41		Los cambios que se presentan en mi trabajo dificultan mi labor
x42		Atiendo clientes o usuarios muy enojados
x43		Atiendo clientes o usuarios muy preocupados
x44		Atiendo clientes o usuarios muy tristes
x45		Mi trabajo me exige atender personas muy enfermas

No.	Dominios	Variable
x46		Mi trabajo me exige atender personas muy necesitadas de ayuda
x47		Atiendo clientes o usuarios que me maltratan
x48		Para hacer mi trabajo debo demostrar sentimientos distintos a los míos
x49		Mi trabajo me exige atender situaciones de violencia
x50		Mi trabajo me exige atender situaciones muy tristes o dolorosas
x51		Mi trabajo me permite desarrollar mis habilidades
x52		Mi trabajo me permite aplicar mis conocimientos
x53		Mi trabajo me permite aprender nuevas cosas
x54		Me asignan el trabajo teniendo en cuenta mis capacidades
x55		Puedo decidir cuánto trabajo hago en el día
x56		Puedo decidir la velocidad a la que trabajo
x57		Puedo cambiar el orden de las actividades en mi trabajo
x58		Los cambios en mi trabajo han sido beneficiosos
x59		Me explican claramente los cambios que ocurren en mi trabajo
x60		Puedo dar sugerencias sobre los cambios que ocurren en mi trabajo
x61	Control sobre el Trabajo	Cuando se presentan cambios en mi trabajo se tienen en cuenta mis ideas y sugerencias
x62		Me informan con claridad cuáles son mis funciones
x63		Me informan cuáles son las decisiones que puedo tomar en mi trabajo
x64		Me explican claramente los resultados que debo lograr en mi trabajo
x65		Me explican claramente el efecto de mi trabajo en la empresa
x66		Me explican claramente los objetivos de mi trabajo
x67		Me informan claramente quien me puede orientar para hacer mi trabajo
x68		Me informan claramente con quien puedo resolver los asuntos de trabajo
x69		La empresa me permite asistir a capacitaciones relacionadas con mi trabajo
x70		Recibo capacitación útil para hacer mi trabajo

No.	Dominios	Variable
x71	Liderazgo y Relaciones Sociales en el Trabajo	Recibo capacitación que me ayuda a hacer mejor mi trabajo
x72		Mi jefe me da instrucciones claras
x73		Mi jefe ayuda a organizar mejor el trabajo
x74		Mi jefe tiene en cuenta mis puntos de vista y opiniones
x75		Mi jefe me anima para hacer mejor mi trabajo
x76		Mi jefe distribuye las tareas de forma que me facilita el trabajo
x77		Mi jefe me comunica a tiempo la información relacionada con el trabajo
x78		La orientación que me da mi jefe me ayuda a hacer mejor el trabajo
x79		Mi jefe me ayuda a progresar en el trabajo
x80		Mi jefe me ayuda a sentirme bien en el trabajo
x81		Mi jefe ayuda a solucionar los problemas que se presentan en el trabajo
x82		Siento que puedo confiar en mi jefe
x83		Mi jefe me escucha cuando tengo problemas de trabajo
x84		Mi jefe me brinda su apoyo cuando lo necesito
x85		Me agrada el ambiente de mi grupo de trabajo
x86		En mi grupo de trabajo me tratan de forma respetuosa
x87		Siento que puedo confiar en mis compañeros de trabajo
x88		Me siento a gusto con mis compañeros de trabajo
x89		En mi grupo de trabajo algunas personas me maltratan
x90		Entre compañeros solucionamos los problemas de forma respetuosa
x91		Hay integración en mi grupo de trabajo
x92		Mi grupo de trabajo es muy unido
x93		Las personas en mi trabajo me hacen sentir parte del grupo
x94		Cuando tenemos que realizar trabajo de grupo los compañeros colaboran
x95		Es fácil poner de acuerdo al grupo para hacer el trabajo
x96		Mis compañeros de trabajo me ayudan cuando tengo dificultades
x97		En mi trabajo las personas nos apoyamos unos a otros
x98		Algunos compañeros de trabajo me escuchan cuando tengo problemas

No.	Dominios	Variable
x99		Me informan sobre lo que hago bien en mi trabajo
x100		Me informan sobre lo que debo mejorar en mi trabajo
x101		La información que recibo sobre mi rendimiento en el trabajo es clara
x102		La forma como evalúan mi trabajo en la empresa me ayuda a mejorar
x103		Me informan a tiempo sobre lo que debo mejorar en el trabajo
x104	Recompensas	En la empresa confían en mi trabajo
x105		En la empresa me pagan a tiempo mi salario
x106		El pago que recibo es el que me ofreció la empresa
x107		El pago que recibo es el que merezco por el trabajo que realizo
x108		En mi trabajo tengo posibilidades de progresar
x109		Las personas que hacen bien el trabajo pueden progresar en la empresa
x110		La empresa se preocupa por el bienestar de los trabajadores
x111		Mi trabajo en la empresa es estable
x112		El trabajo que hago me hace sentir bien
x113		Siento orgullo de trabajar en esta empresa
x114		Hablo bien de la empresa con otras personas

Tabla 2 - 3. Variables musculo esqueléticas evaluadas. Tomado de: (Sauter & Murphy, 1984).

Datos musculo esqueléticos	Variables
Dolor de cabeza constante	M1
Migraña	M2
Dolor de hombros	M3
Dolor en los brazos	M4
Dolor de espalda – Alta –Media o Baja	M5

Tabla 2 - 4. Variables fisiológicas evaluadas. Tomado de: (Sauter & Murphy, 1984).

Datos fisiológicos tomados en pacientes que lo necesitaron y participaron voluntariamente	Variables
HR	Frecuencia cardíaca
EDA	Actividad eléctrica de la piel
Emg	Electromiografía
Clases	4 = Riesgo Muy Alto 3 = Riesgo Alto 2 = Riesgo Medio 1 = Riesgo Bajo 0 = Riesgo Muy Bajo

IV. Capítulo 3 - Materiales y Métodos

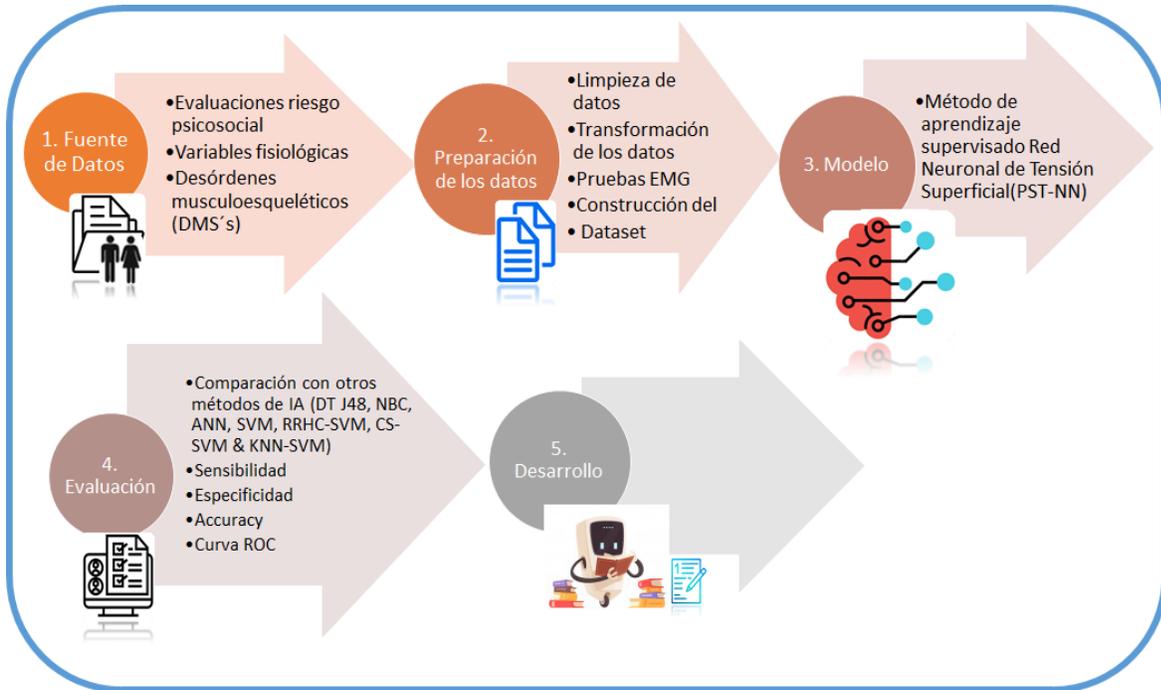
El objetivo del presente capítulo es diseñar una metodología de clasificación basado en técnicas inteligentes para identificar el grado de riesgo psicosocial en docentes de colegios públicos de Colombia que permita minimizar tiempos de proceso en la evaluación y costos de recursos, facilitando la tarea para el psicólogo laboral, con el fin de contribuir a la toma de decisiones organizacionales y de Seguridad y Salud en el Trabajo en el sector educativo público del país. La novedad consistió en incorporar el modelo de la tensión física superficial al de redes neuronales artificiales emulando las condiciones físicas de la tensión de superficie en algunos líquidos para poder clasificar las variables con las que finalmente se puede identificar el nivel de riesgo psicosocial.

3.1 Factores de riesgo psicosocial y proceso de construcción del modelo

En el presente documento, se destacan cinco aspectos, que afectan en gran medida la efectividad de los modelos de predicción en la ocurrencia de riesgo psicosocial en docentes de colegios públicos (**Figura 3 - 1**).

La fuente de datos se refiere a los datos recopilados a lo largo de las evaluaciones de riesgo psicosocial realizadas en la población docente de la ciudad y su área metropolitana, así como a los datos fisiológicos y los diagnósticos musculoesqueléticos obtenidos.

Figura 3 - 1. Pasos para la construcción del modelo para la identificación del nivel de riesgo psicosocial en docentes de colegios públicos.



Fuente: Elaboración propia.

2) La preparación de datos se refiere al proceso mediante el cual los datos se ajustan al modelo de predicción. Se relacionan variables que están relacionados con la gravedad del riesgo psicosocial en los entornos de trabajo académico público. Estos se aplican para su uso en el modelo de predicción, con el fin de identificar, comprobar y aumentar su rendimiento.

3) El modelado es el proceso de diseño, construcción y funcionamiento del método desarrollado, del ajuste al tipo de datos y su exactitud predictiva.

4) La evaluación es el proceso de evaluación y comparación de la efectividad de los modelos matemáticos y su exactitud predictiva. Las diferentes mediciones generalmente permiten evaluar la eficiencia de cualquier modelo de predicción.

5) El desarrollo se refiere a la predicción del nivel de riesgo y los resultados a perseguir, según lo dispuesto por el modelo de predicción, basado en los planes de prevención formulados y descritos en la literatura científica.

3.2. Base de datos

3.2.1 Factores de riesgo psicosocial, variables fisiológicas y musculoesqueléticas

Las variables consideradas para el diseño del modelo de Red Neuronal de Tensión superficial se enumeran en las **Tabla 3 - 1– Tabla 3 - 3**.

3.3. Pre procesamiento de los datos

3.3.1. Conjunto de datos

En relación con los factores de riesgo psicosocial, estos pueden clasificarse en dos clases principales: los que tienen efectos negativos en la salud y los que contribuyen positivamente al bienestar del trabajador. Aunque ambos están presentes en todos los entornos de trabajo, el presente estudio considera aquellos que afectan negativamente la salud en las organizaciones académicas de las escuelas públicas (Bruhn & Frick, 2011; Dediú et al., 2018; El-Batawi, 1988; Lippel & Quinlan, 2011; Weissbrodt & Giaunque, 2017).

En este estudio, se analizaron los resultados de las evaluaciones para medir el nivel de riesgo psicosocial de 5443 docentes, en cinco ciudades de Colombia. Las cinco ciudades colombianas estaban en la misma área geográfica (Cali, Palmira, Yumbo, Candelaria y Jamundí). Los datos se obtuvieron de una base de datos inicial ya construida y se evaluaron otros docentes durante un período de un año y ocho meses para complementar la información. El protocolo para la recolección de los datos es la que indica la Resolución 2646 de 2008 legal vigente, incluido el consentimiento informado (Anexo E).

Tabla 3 - 1. Variables para factores de riesgo intralaboral. Tomado de: (Villalobos et al., 2010).

Variables riesgo psicosocial		
Factor	Dominio	Descripción
Liderazgo y relaciones sociales en el trabajo (L)	Características Liderazgo(L ₁)	Atributos de la gestión de los jefes inmediatos en relación con la planificación y asignación del trabajo, consecución de resultados, resolución de conflictos, participación, motivación, apoyo, interacción y comunicación con sus colaboradores.
	Retro alimentación desempeño(L ₂)	Información que un trabajador recibe sobre la forma como realiza su trabajo. Esta información le permite identificar sus fortalezas y debilidades y tomar acciones para mantener o mejorar su desempeño.
Control (C)	Claridad del rol(C ₁)	Definición y comunicación del papel que se espera que el trabajador desempeñe en la organización, específicamente en torno a los objetivos del trabajo, las funciones y resultados, el margen de autonomía y el impacto del ejercicio del cargo en la empresa.
	Capacitación(C ₂)	Actividades de inducción, entrenamiento y formación que la organización brinda al trabajador con el fin de desarrollar y fortalecer sus conocimientos y habilidades.
	Uso de habilidades y destrezas y desarrollo(C ₃)	Posibilidad que el trabajo le brinda al individuo de aplicar, aprender y desarrollar sus habilidades y conocimientos.
Demandas del trabajo (D)	Demandas ambientales y de esfuerzo físico(D ₁)	Son condiciones de tipo físico (ruido, iluminación, temperatura, ventilación), químico, biológico (virus, bacterias, hongos o animales), de diseño del puesto de trabajo, de saneamiento (orden y aseo), de carga física y de seguridad industrial.

Dominio		Descripción
	Demandas emocionales(D_2)	Exigencias emocionales demandan del trabajador habilidad para: a) entender las situaciones y sentimientos de otras personas y b) ejercer autocontrol de las emociones o sentimientos propios con el fin de no afectar el desempeño de la labor.
	Demandas cuantitativas(D_3)	Exigencias relativas a la cantidad de trabajo que se debe ejecutar, en relación con el tiempo disponible para hacerlo.
	Influencia del ambiente laboral sobre el extra laboral (D_4)	Exigencias de tiempo y esfuerzo que se hacen a un individuo en su trabajo, impactan su vida extra laboral (Familiar).
	Demandas de carga mental (D_5)	Se refieren a las demandas de procesamiento cognitivo que implica la tarea y que involucran procesos mentales superiores de atención, memoria y análisis de información para generar una respuesta.
	Demandas de la jornada de trabajo(D_6)	Exigencias del tiempo laboral que se hacen al individuo en términos de la duración y el horario de la jornada, así como de los periodos destinados a pausas y descansos periódicos.
Recompensas (R)	Reconocimiento y compensación(R_1)	Conjunto de retribuciones que la organización le otorga al trabajador en contra prestación al esfuerzo realizado en el trabajo. Estas retribuciones corresponden a reconocimiento, remuneración económica, acceso a los servicios de bienestar y posibilidades de desarrollo.

Tabla 3 - 2. Variables fisiológicas. Tomado de: (Sauter & Murphy, 1984).

Variables fisiológicas	
Frecuencia Cardíaca (P_1)	La frecuencia cardíaca es la velocidad del latido del corazón medida por la cantidad de contracciones (latidos) del corazón por minuto (lpm).
Actividad eléctrica de la piel (P_2)	Propiedad del cuerpo humano que causa variación continua en las características eléctricas de la piel. La conductancia de la piel puede ser una medida de las respuestas emocionales y comprensivas.
Electromiografía (P_3)	Es una técnica de medicina electro-diagnóstica para evaluar y registrar la actividad eléctrica producida por los músculos.

Tabla 3 - 3. Síntomas musculoesqueléticos.

Variables físicas (Relativas al ausentismo laboral y factores psicosociales)	
Síntomas	Descripción
Dolor de cabeza y dolor cervical (M_1)	Un dolor de cabeza en general es un signo de estrés o angustia emocional, y puede asociarse con migraña o presión arterial alta, ansiedad o depresión. Algunos pacientes experimentan dolor de cabeza durante 2 horas o menos. (Comité de Clasificación de Dolor de Cabeza de la International Headache Society (IHS), 2013).
Migraña (M_2)	Las migrañas pueden asociarse con un dolor de cabeza intenso que a menudo aparece en un lado de la cabeza. Tienden a afectar a personas de 15 a 55 años. Los síntomas incluyen hiperactividad, hipoactividad, depresión, fatiga y rigidez en el cuello y / o dolor intenso (Comité de Clasificación de Dolor de Cabeza de la Sociedad Internacional de Dolor de Cabeza (IHS), 2013).
Dolor de hombros (M_3)	El dolor es provocado o agravado por el movimiento. El dolor y la rigidez generalmente restringen el uso de las extremidades superiores y, por lo tanto, limitan las actividades diarias durante el trabajo (Van der Heijden, 1999).
Dolor en los brazos (M_4)	El dolor en el brazo es causado por movimientos repetitivos en el trabajo, por lo general, los síntomas se describen como dolor, incomodidad o rigidez que ocurre desde los hombros hasta los dedos.
Dolor de espalda (M_5)	El dolor de espalda en el trabajo generalmente puede afectar a personas de cualquier edad, levantar objetos pesados, movimientos repetitivos y sentarse en un escritorio todo el día pueden producir una lesión.

El conjunto de datos se puede consultar en:

<https://zenodo.org/record/1298610> (Mosquera, 2019; Mosquera, 2018).

El conjunto de datos contiene información sobre las siguientes variables: i) sociodemográficas (sexo, edad, estado civil y nivel de ingresos); ii) psicosocial; iii) fisiológica, y; iv) variables asociadas con dolor y trastornos musculoesqueléticos. En total, 115 variables fueron inicialmente incluidas en este estudio.

En general, el conjunto de datos se obtuvo de manera presencial sobre los docentes aplicando la evaluación de la batería de riesgos psicosocial, los docentes que habían manifestado ausentismo laboral por enfermedad general se descartaron para realizar la evaluación de variables físicas y fisiológicas y los docentes que manifestaron ausentismo laboral por enfermedad profesional, se tuvieron en cuenta para revisar su historia clínica, aplicar las pruebas de tipo físico y fisiológico. Estos docentes que participaron en el muestreo, lo hicieron de forma voluntaria.

3.3.2. Análisis de datos

- Validación de problemas osteomusculares

Para probar los problemas musculoesqueléticos declarados en pacientes con un cierto nivel de riesgo (medio, alto y muy alto) y confirmados en su historia clínica, se realizó una electromiografía de superficie para corroborar su diagnóstico. Se realizó en una muestra compuesta por 495 pacientes que decidieron participar voluntariamente (**Figura 3 - 2**), la toma de datos la realizó un tecnólogo en electrónica que participó en el proyecto y la adquisición de la señal se hizo con equipos BITalino (Bitalino, 2017), los resultados fueron revisados por un médico que colaboró en el proyecto, para identificar si realmente presentaban problemas osteomusculares (Anexo E).

Figura 3 - 2. Aplicación de la batería de riesgos intralaboral a cada docente y evaluación de condiciones fisiológicas a docentes con problemas médicos declarados.



Fuente: Elaboración propia.

- Variables de entrada y salida.

Entre las variables de factores de riesgo asociadas con el análisis del entorno laboral, hay un total de 114 variables de entrada: $X_{ij} = (\text{factores psicossociales}, j = 1, \dots, 123)$, $P_{ij} = (\text{factores fisiológicos}, j = 1, \dots, 3)$ y $M_{ij} = (\text{síntomas musculoesqueléticos}, j = 1, \dots, 5)$, donde, i es el sujeto en estudio. Posteriormente, las variables redundantes (X_{ij}) se filtraron mediante el uso de la importancia de rango de los predictores utilizando el procedimiento del algoritmo ReliefF (11) (Robnik-S̃ & Kononenko, 2003), con el objetivo de identificar variables generadoras de ruido en el conjunto de datos utilizando los criterios métricos de Chebyshev.

El objetivo del procedimiento del algoritmo de ReliefF, que localiza predictores importantes en los 10 vecinos más cercanos, es agrupar las 123 variables independientes de X_{ij} . Los números de predictores se enumeran por clasificación, y el algoritmo selecciona los predictores de mayor importancia. Los pesos producen valores de peso en el mismo orden que el de los predictores. Las distancias entre pares de variables, en este peso, se miden una vez más, y se elige la variable con el valor total más bajo (distancia), que produce 12 variables X_{ij} por grupo. Además, agrega tres variables P_{ij} fisiológicos ($j = 1, \dots, 3$) y cinco variables de síntomas musculoesqueléticos $M_{ij} = (j = 1, \dots, 5)$, como se muestra en las

Tabla 2 - 2 - Tabla 2 - 4.

El algoritmo reconoce las variables con el valor más bajo y castiga esos predictores (riesgo asociado con cada X_{ir} individual) que produce diferentes valores para los vecinos en el mismo grupo (grupo de factores de riesgo F_{ij}), y aumenta aquellos que producen diferentes valores para los vecinos en diferentes grupos. ReliefF inicialmente define los pesos del predictor R_{ij} en 0, y el algoritmo luego selecciona un valor aleatorio X_{ir} , iterativamente. Se identifican los valores k más cercanos X_{ir} para cada grupo, y se actualizan todos los pesos predictores F_{ij} para los vecinos más cercanos X_{iq} (**Figura 3 - 3**), (Robnik-S̃ & Kononenko, 2003, p. 26).

$$W[A] := W[A] - \sum_{j=1}^k \frac{\text{diff}(A, R_i, H_j)}{m \cdot k} + \sum C \neq \text{class}(R_i) + \left[\frac{P(C)}{1 - P(\text{class}(R_i))} \sum_{j=1}^k \frac{\text{diff}(A, R_i, M_j(C))}{m \cdot k} \right] \quad (12)$$

Dónde,

R_i es el número de valores seleccionado aleatoriamente.

H_i es k vecinos cercanos clasificados correctamente (k -nn con la misma clase).

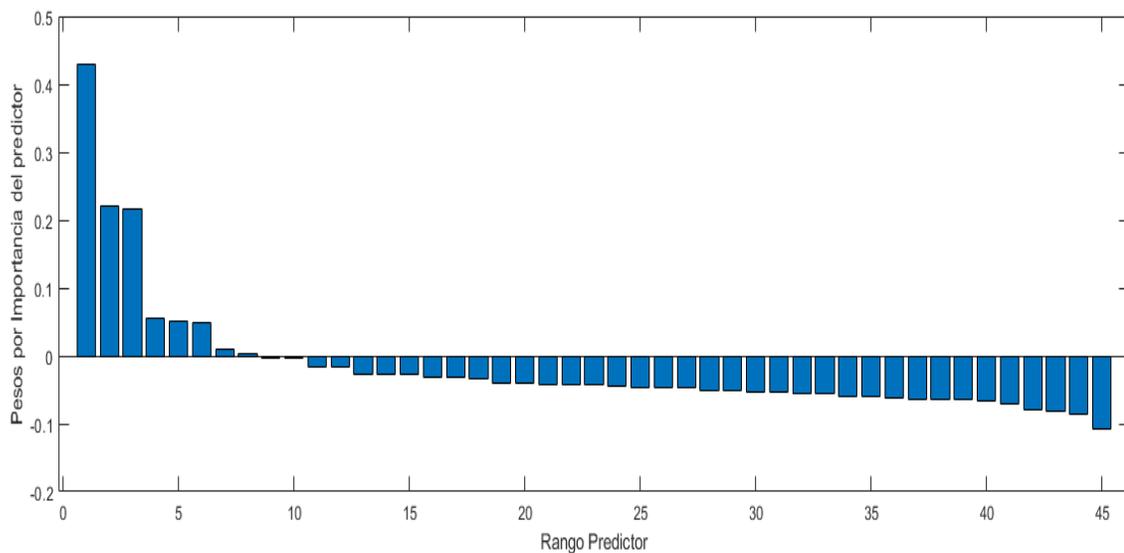
$M_j(C)$ es k vecinos no correctamente clasificados (k -nn con diferente clase).

$W[A]$ es la calidad de la estimación para todos los atributos A para R_i , y H_j y valores mal clasificados $M_j(C)$.

$1 - P(class(R_i))$ es la suma de las probabilidades para las clases mal clasificadas.

m es el tiempo de proceso repetido.

Figura 3 - 3. Aplicación algoritmo ReliefF para reducción de la dimensionalidad de los datos.



Fuente: Elaboración propia.

Finalmente, acorde a sugerencias por parte de psicólogos laborales se decide:

En total, seleccionar 20 variables de entrada $E_{ij} = X_{ij} + P_{ij} + M_{ij}$: doce variables $X_{ij} = (j = 1, \dots, 12)$, que constituían variables psicosociales; tres variables fisiológicas $P_{ij} = (j = 1, \dots, 3)$, y; cinco variables asociadas con los síntomas musculoesqueléticos $M_{ij} = (j = 1, \dots, 5)$, como

se muestra en las **Error! Reference source not found. – Error! Reference source not found.**

Las 20 variables de entrada se normalizaron, de acuerdo con la Ecuación 13.

$$E_{normalized} = \frac{(E - E_{min})}{E_{max} - E_{min}} \quad (13)$$

Donde E corresponde a la variable a normalizar, E_{max} es el valor máximo de cada variable, E_{min} es el valor mínimo y $E_{normalized}$ es la variable normalizada dentro del rango -1 a 1. Las variables se identifican como las variables que afectan a la persona desde $E_{1...}, E_{20}$.

3.4. Desarrollo de algoritmos previos a la obtención del modelo propuesto

Los modelos de optimización o algoritmos de optimización como random restart hill climbing (El Yafrani & Ahiod, 2018; Kato et al., 2018) y cuckoo search (Cheng et al., 2019; Li et al., 2021; Meneses et al., 2020) han sido utilizados en gran variedad de problemas para lograr hallar soluciones en espacios de búsqueda amplia, buscando maximizar la eficiencia de los resultados, estos, junto a algoritmos de clasificación lineal como, las máquinas de soporte vectorial han ayudado a mejorar la función de clasificación. Algunos autores han trabajado las máquinas de soporte vectorial para clasificar estrés y características relacionadas con el estrés como rasgos característicos de la personalidad (Hadi et al., 2019; Lotfan et al., 2019; Priya et al., 2020; Rodríguez-Arce et al., 2020; Sriramprakash et al., 2017; Zubair & Yoon, 2020).

Antes de llegar a la metodología diseñada y finalmente propuesta, se realizarán aproximaciones con algunos algoritmos híbridos (Mosquera, Gómez, et al., 2018), luego, se realizaron dos aproximaciones con algoritmos híbridos compuestos por algoritmos de optimización y de un clasificador de margen máxima con kernel polinomial (SVM) que abordaremos de forma general para luego centrarnos en el modelo propuesto finalmente. Dichos desarrollos están en fase de publicación en revistas indexadas.

3.4.1. Algoritmo de ascenso de colinas con reinicio aleatorio – Máquinas de Soporte Vectorial (RRHC-SVM)

El algoritmo de ascenso de colinas con reinicio aleatorio (RRHC) selecciona los valores iniciales, con los que se formula una solución inicial de forma aleatoria. RRHC ofrece valores optimizados, que se toman para los parámetros de entrada a la máquina de soporte vectorial (SVM), con el fin de ejecutar soluciones iniciales para modelar el sistema. Cuando se combinan los dos modelos, se elimina el problema de valores inadecuados de los parámetros de entrada de SVM y se eligen los datos de entrada óptimos para la clasificación de datos del modelo. Cuando el RRHC busca parámetros de entrada, los valores se dan aleatoriamente con intervalos específicos. Después, el algoritmo RRHC se usa repetidamente y genera reinicios aleatorios para buscar los valores de entrada óptimos del modelo de predicción de SVM, considerando que los valores anteriores no se repiten. Finalmente, finaliza cuando el modelo de predicción maximiza el valor de clasificación con los parámetros de entrada optimizados.

Aquí, aplicamos $n = 10$ objetos asignados a una ubicación, el flujo entre estos objetos se da en una matriz N , y la distancia entre cada dos ubicaciones en una matriz M . La solución está representada por la permutación p de n números. El objetivo es encontrar la asignación p^* que minimice la suma de los productos de distancia y flujo entre cada dos parámetros:

$$p^* = \text{riskmin}_p \sum_{i=1}^n \sum_{j=1}^n F_{i,j} D_{p_i,p_j} \quad (14)$$

Dónde,

p_i representa la ubicación del objeto i .

Y el espacio de la solución es, un individuo i que se elige al azar (al comienzo de cada simulación) y a cada vecindario $j = 1, \dots, n$ y a cada individuo i se le asigna una ubicación j : $p_i = j$. El algoritmo anterior se aplica con $\epsilon = 0.1$. Para cada iteración con búsqueda aleatoria.

3.4.2. Algoritmo de búsqueda cuckoo – Máquinas de Soporte Vectorial (CS-SVM)

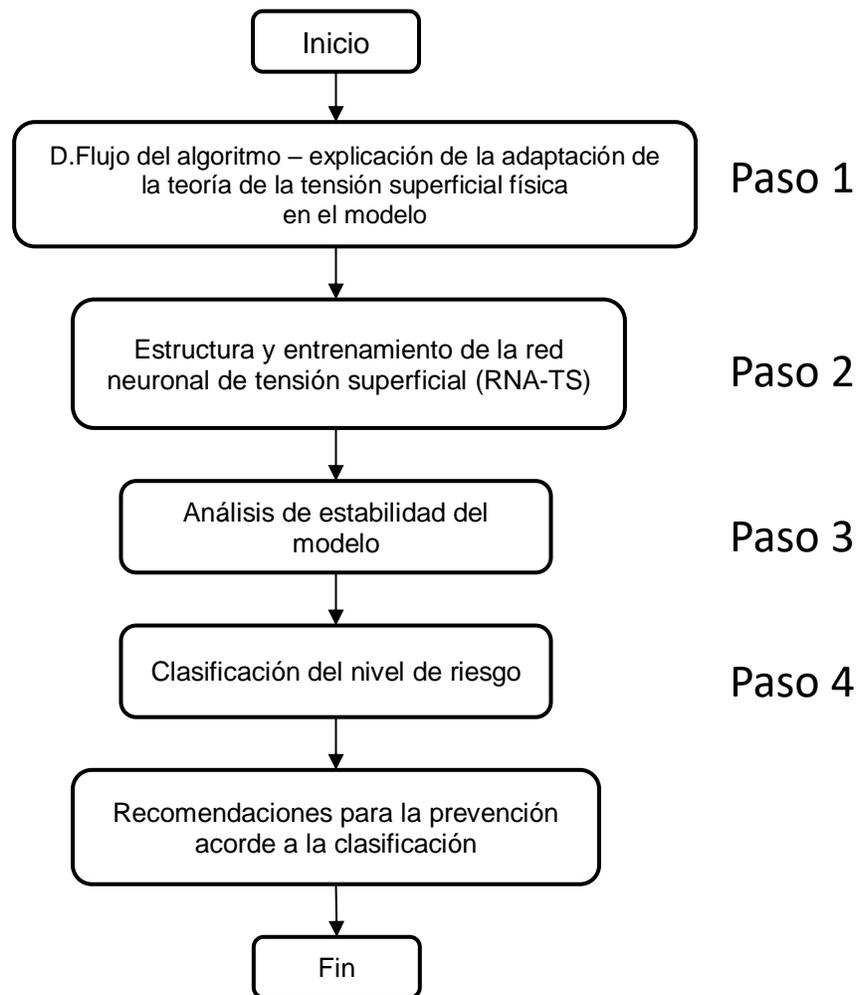
El proceso de búsqueda cuckoo, comienza con la generación de una población inicial aleatoria a partir de los nidos huéspedes (riesgo individual i y su variable de factor de riesgo asociado j). El algoritmo itera hasta que encuentra los criterios de detención. Se crea una nueva solución mediante una función de desplazamiento. Esta nueva solución se compara aleatoriamente con otra, y si la primera es mejor, reemplaza a la segunda. Los peores nidos se abandonan, dependiendo de la probabilidad de P_a dada. En este caso, la probabilidad de P_a fue de 0.1%, hubo cinco nidos iniciales y 2.000 generaciones, según la experiencia.

De forma general, primero viene el número de nidos (el espacio de búsqueda de los valores que pertenecen a las variables de riesgo, ($i = V.Risk_1, V.Risk_2 \dots V.Risk_{20}$), seguido de los porcentajes de los peores (el valor del menor riesgo), que corresponde al porcentaje de los peores nidos P_a . Estos se abandonan en cada iteración. Por último, viene el número de generaciones, o número máximo de iteraciones, cada iteración se almacena y se procesa en el clasificador con los valores optimizados para su clasificación hasta obtener el mejor número de entradas de clasificación, así:

$$"x = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \dots \\ x_{20} \end{pmatrix} \rightarrow w = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ \dots \\ w_{20} \end{pmatrix} "$$
(15)

La metodología diseñada se compone de 4 pasos como se presenta en la **Figura 3 - 4**, los cuales se explican a continuación:

Figura 3 - 4. Flujograma de la metodología propuesta.



Fuente: Elaboración propia.

3.5. Paso 1. El algoritmo propuesto de Red Neuronal de tensión superficial (RNA-TS)

Suárez Sánchez et al., (2011) trabajaron en el desarrollo de un modelo SVM capaz de clasificar correctamente a los trabajadores en dos grupos, los que han sufrido un accidente laboral en los últimos doce meses y los que no, según sus respuestas a la Sexta Encuesta Nacional de Condiciones del Trabajo y sin tener en cuenta la variable en la que se dice si han sufrido un accidente o no. Por lo tanto, el modelo obtenido lo considerarían como un modelo predictivo. Consideraban que SVM era un clasificador lineal general que era una extensión del Perceptrón. El objetivo era lograr identificar las variables primordiales asociadas a accidentalidad laboral que ocasionaban accidentes laborales, entre ellas, los factores psicosociales, obteniendo una aproximación en la clasificación. El modelo de clasificación de máquinas de soporte vectorial, básicamente clasificaba a los individuos con unas etiquetas de clase de haber tenido un accidente (1) o no haber tenido un accidente (0). El modelo proponía lo interesante que era la influencia de las variables psicosociales en la ocurrencia de accidentes laborales, según la encuesta. Basado en el desarrollo y en la importancia de mejorar las condiciones laborales se decide diseñar y desarrollar el modelo que trataremos en éste capítulo.

El enfoque se basa en la teoría de la tensión superficial de los líquidos (Jasper, 1972; Macleod, 1923; Tyson & Miller, 1977) dado en la ecuación (16). La teoría de la tensión superficial aborda la cohesión entre las moléculas en un líquido y su relación energética con el exterior, generalmente un gas. La superficie de un líquido, cuando se somete a una fuerza que rompe el estado energético de cohesión molecular, hace que el objeto que produce la fuerza interna en el líquido se hunda (

Figura 3 - 5).

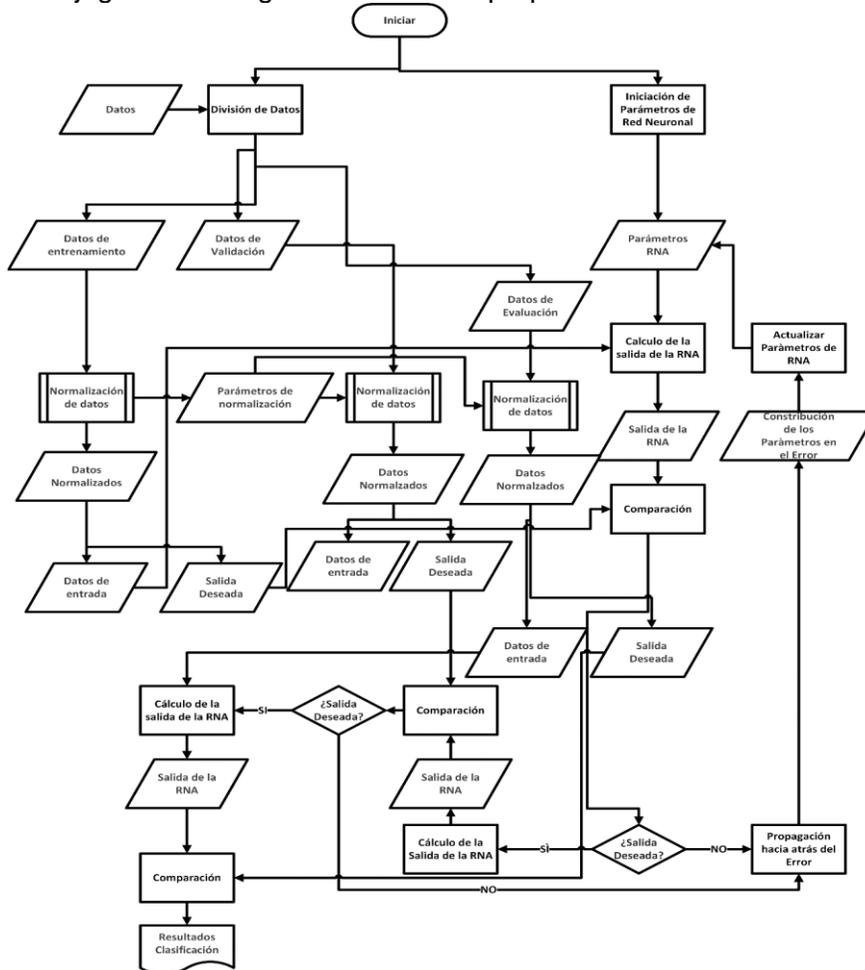
En esta sección, se propone un modelo híbrido y estandarizado basado en la tensión superficial física para predecir los niveles de riesgo psicosocial, con el objetivo de guiar el desarrollo futuro del área de investigación denominándola psychosocial analytics (**Figura 3 - 6).**

Figura 3 - 5. Strider de agua sobre el líquido debido a la tensión superficial.



Fuente: Adaptado de Waterstrider [Fotografía], por (Wikimedia Commons, 2006), (<https://commons.wikimedia.org/wiki/File:WaterstriderEnWiki.jpg>). CC BY 3.0

Figura 3 - 6. Flujoograma del algoritmo RNA -TS propuesto.



Fuente: Elaboración propia.

Esta propuesta busca emular la teoría de la tensión superficial, desde el punto de vista del análisis psicosocial, en relación con los factores de riesgo presentes en los entornos de trabajo y sus grados de riesgo desde el punto de vista de la mejora del modelo de aprendizaje automático. Utilizará y adaptará dicha teoría para mejorar la clasificación del riesgo y modificar los parámetros necesarios (el número de capas, nodos, pesos y umbrales) para reducir el tiempo de proceso en la evaluación, la dimensionalidad de los datos y aumentar la exactitud. Este nuevo método clasificará los datos de entrada de la red, por lo que hay una capa de entrada con veinte nodos de entrada, dos capas ocultas de 80 neuronas, una capa de tensión superficial oculta con dos nodos y una salida de cuatro nodos, para clasificar el riesgo.

$$\gamma = \frac{F}{2L} \quad (16)$$

Dónde,

γ es la tensión superficial, es la dimensión de la fuerza por unidad de longitud, o de energía por unidad de área (en el modelo γ es el nivel de riesgo o grado de riesgo, la clasificación), F es la fuerza requerida para lograr que uno de los lados comience a deslizarse, lo que significaría la presión psicológica ejercida sobre el individuo por los factores a los cuales se encuentra expuesto en el ambiente laboral, L la longitud del lado móvil, es decir, la capacidad del individuo de afrontar la o no las situaciones que lo afectan psicosocialmente, y, la razón del 1/2 es que la película tiene dos superficies, la superior y la inferior (Macleod, 1923), en este caso, el medio, correspondería al ambiente laboral al cual pertenece el docente, se podría entender como las condiciones laborales del ambiente de trabajo donde se encuentra expuesto el docente.

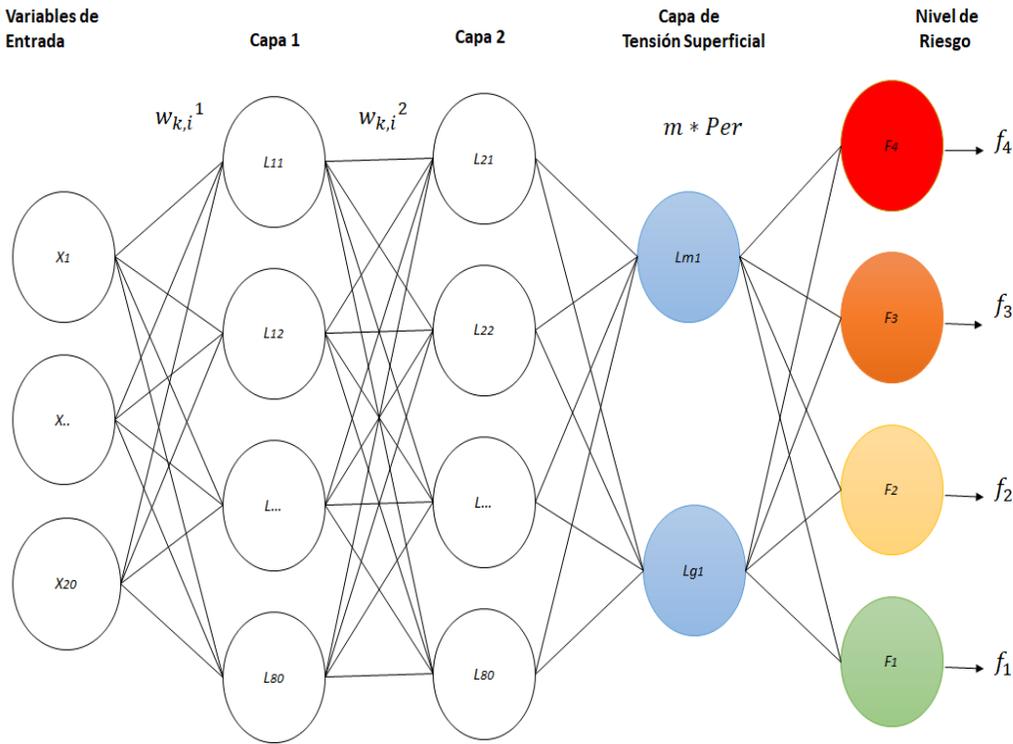
En este modelo, la multiplicación del perímetro de un objeto por la tensión superficial de un líquido produce la fuerza que ejerce un líquido sobre su superficie, sobre el objeto, para evitar que dicha tensión se rompa. Como tal, si el peso de un objeto es mayor que la fuerza ejercida por el líquido sobre su superficie, el objeto tenderá a hundirse (lo que equivaldría a la afectación que sufre el docente debido a los factores de riesgo psicosociales a los cuales se encuentra expuesto).

Dicha teoría presenta una ventaja importante, dada la reducción de la dimensionalidad de la variable de entrada no lineal, considerando la variable dependiente categórica, que permite reducir 20 variables a 2 variables para llevar a cabo el proceso de clasificación.

3.5.1. Conversión de variables matemáticas a variables físicas

Las 20 variables de entrada E_{ij} de los datos de entrenamiento se convierten en dos variables físicas, perímetro (Per) y masa (m), a lo largo de una red neuronal artificial con cuatro capas. Tres de estos constituyen la arquitectura de una red neuronal estándar, con la diferencia de que el último nivel contiene un nuevo modelo de red neuronal basado en la tensión superficial física (Adamson & Gast, 1967).

Se utilizaron ochenta neuronas en las capas uno y dos, debido al hecho de que a medida que aumentaba este número, no se registraron cambios sustanciales. Además, solo se usaron dos neuronas para la capa 3, para anexar la nueva capa de tensión superficial propuesta. La arquitectura de la red de clasificación neuronal artificial se muestra en la **Figura 3 - 7**. Esto incluye las tres capas neuronales estándar, así como una capa con un diseño novedoso.

Figura 3 - 7. Red Neuronal de tensión superficial física (RNA-TS).

Fuente: Elaboración propia.

3.5.2. Inicialización de parámetros de red neuronal

Para la inicialización de los parámetros de la red neuronal, se utiliza el algoritmo Nguyen-Widrom (Andayani et al., 2017; Pavelka & Prochazka, 2004), en el que se generan parámetros aleatorios. Sin embargo, la ventaja de esto es que los parámetros distribuyen las regiones neuronales activas de manera mucho más uniforme en capas, lo que mejora el entrenamiento de la red neuronal, ya que presenta un error de velocidad de aprendizaje más bajo desde el principio.

3.6. Paso 2. Entrenamiento de la red neuronal artificial (PST-NN)

3.6.1. Cálculo de salida de la capa 1

Las 20 variables de entrada de un individuo específico del conjunto de entrenamiento, posee un vector que se llamará E , que pasan a través de una capa inicial de 80 neuronas. Cada neurona tiene 20 parámetros, llamados pesos, que multiplican cada variable de entrada del vector E . A esta multiplicación se agrega un parámetro llamado umbral de la función b o sesgo (bias b). Que está asociado con cada neurona, lo que resulta en la salida parcial de la capa 1. Este procedimiento puede describirse a lo largo de la ecuación (17).

$$y_k^1 = \left(\sum_{i=1}^{20} E_i * w_{k,i}^1 \right) + b_k^1 \text{ for } k = 1 \text{ to } 80 \quad (17)$$

$$y^1 = \{y_1^1, y_2^1 \dots y_{80}^1\} \quad (18)$$

Donde E_i es la variable i del individuo elegido del conjunto de entrenamiento, $w_{k,i}^1$ es el peso de la neurona k en la capa 1, que se multiplica por la variable i , b_k^1 es el sesgo de la función de la neurona k en la capa 1, que se suma al total, y y_k^1 es el resultado de cada k neurona. Estos 80 resultados pueden estar representados por el vector y^1 , y y^1 pasa a través de una función de transferencia de tangente hiperbólica, ya que esta es una función de transferencia continua y se recomienda para procesos de reconocimiento de patrones.

La salida de la capa 1 se describe en la ecuación (19):

$$Y_k^1 = \frac{2}{1 + e^{-2*y_k^1}} - 1 \text{ for } k = 1 \text{ to } 80 \quad (19)$$

$$Y^1 = \{Y_1^1, Y_2^1 \dots Y_{80}^1\} \quad (20)$$

Donde e es la función exponencial y Y^1 es la salida final para la Capa 1 y está compuesta por 80 salidas: una para cada neurona.

3.6.2. Cálculo de salida de la capa 2

Las 80 salidas de la Capa 1, Y^1 , se convierten en las entradas para la Capa 2, que presenta el mismo número de neuronas que la Capa 1. Como tal, de acuerdo con el procedimiento realizado en la Capa 1, se obtienen las ecuaciones (21) y (22):

$$y_k^2 = \left(\sum_{i=1}^{80} Y_i^1 * w_{k,i}^2 \right) + b_k^2 \text{ for } k = 1 \text{ to } 80 \quad (21)$$

$$y^2 = \{y_1^2, y_2^2 \dots y_{80}^2\} \quad (22)$$

Donde Y_i^1 es la salida de la neurona i de la capa 1, $w_{k,i}^2$ es el peso de la neurona k , asociado con la salida de la neurona i en la capa 1, b_k^2 es el sesgo de la neurona k en la capa 2, y y^2 incluye las 80 respuestas de cada neurona, antes de pasar por la función de transferencia. Para obtener la salida final de la capa 2 (Y^2), se aplica la función de transferencia hiperbólica (23):

$$Y_k^2 = \frac{2}{1 + e^{-2*y_k^2}} - 1 \text{ for } k = 1 \text{ to } 80 \quad (23)$$

$$Y^2 = \{Y_1^2, Y_2^2 \dots Y_{80}^2\} \quad (24)$$

3.6.3. Cálculo de salida de la capa 3

Las 80 salidas para la capa 2 se ingresan en la capa 3, que contiene dos neuronas (25) y (26):

$$Y_k^3 = \left(\sum_{i=1}^{80} Y_i^2 * w_{k,i}^3 \right) + b_k^3 \text{ for } k = 1 \text{ to } 2 \quad (25)$$

$$Y^3 = \{Y_1^3, Y_2^3\} \quad (26)$$

Donde Y_i^2 es la salida de la neurona i en la Capa 2, $w_{k,i}^3$ es el peso de la neurona k en la Capa 3, que multiplica la salida de la neurona i en la Capa 2, y Y_k^3 es la salida final de cada una de las dos neuronas representadas en el vector Y^3 .

Estas dos variables de salida se consideran masa y perímetro, respectivamente. En términos de masa (m), y perímetro (Per), estos dos parámetros de ecuación de tensión superficial intervienen para reemplazar la función sigmoidea de la red y pueden reducir la variabilidad y la dimensionalidad de los datos para mejorar la respuesta en la clasificación y el error resultante. Como tal, la función de transferencia no se aplica, para evitar la reducción del rango de respuesta:

$$Y^3 = \{Y_1^3, Y_2^3\} = \{m, Per\} \quad (27)$$

3.6.4. Cálculo de salida de la capa de tensión superficial

Las dos salidas de capa 3, m y Per (masa y perímetro), contenido en Y^3 , entran en una capa final llamada la capa de tensión superficial. Esta se compone de cuatro neuronas (para el nivel de riesgo). Cada uno de estos contribuye a un equilibrio de fuerzas definido por la siguiente ecuación (28), (29):

$$O_k = 1 - e^{\frac{F}{2L}} \text{ for } k = 1 \text{ to } 4 \quad (28)$$

Dónde,

$$O_k = 1 - e^{\frac{-m*g}{T_k*Per}} \text{ for } k = 1 \text{ to } 4 \quad (29)$$

$$O = \{O_1, O_2, O_3, O_4\} \quad (30)$$

Con

$$T_k = \{22.1; 47.7; 72.8; 425.41\} \quad (31)$$

Donde m es la masa que corresponde a la salida de la primera neurona de la capa 3, g es el valor de la constante de gravedad $9.8 \frac{m}{s^2}$. La multiplicación de la masa por la gravedad $m * g$ produce el peso de un objeto. Per , el perímetro, es la salida de la segunda neurona, de la capa 3, y T_k es el valor de la tensión superficial en la neurona k .

La multiplicación del perímetro de un objeto por la tensión superficial de un líquido da como resultado la fuerza ejercida por un líquido sobre su superficie sobre un objeto para no dejarse romper. Por lo tanto, si el peso de un objeto es mayor que la fuerza ejercida por el líquido sobre la superficie, tenderá a hundirse, lo que dará como resultado el nivel de clasificación.

Es decir, los factores de riesgo psicosociales intralaborales presentes en el ambiente laboral presionan sobre el desempeño del trabajador, el cual presenta una resistencia individual ante los factores para soportar y continuar su vida laboral normal, cuando estos factores sobrepasan la capacidad individual, el docente se verá afectado de alguna manera, lo que ocasionará que tenga algún grado de riesgo psicosocial.

Y, T es el grupo de cuatro tensiones superficiales de cuatro líquidos, como se incluye en la **Tabla 3-4**.

Tabla 3 - 4. Tensión superficial de líquidos (Jasper, 1972).

Líquido	Valor tension superficial (<i>mN/M</i>) at 20°C
Etanol	22.1
Ethylene glycol	47.7
Agua	72.8
Mercurio	425.41

Se utilizaron los cuatro líquidos mostrados anteriormente, ya que son comunes, relativamente conocidos y presentan diferentes tensiones superficiales. Aquí, la idea principal es la relación que existe entre las cuatro tensiones superficiales y los diferentes pesos de los objetos que pueden romper la tensión superficial del líquido.

Para nuestro modelo, la tensión superficial de cada líquido es similar a cada nivel de riesgo psicosocial, donde el nivel de riesgo más bajo será la tensión superficial del etanol, y el nivel de riesgo muy alto es equivalente a la tensión superficial del mercurio. En este sentido, cuando una persona tiene, según la evaluación psicosocial, un alto nivel de riesgo, los parámetros en la nueva neurona de tensión superficial serán equivalentes a haber recorrido la tensión superficial del alcohol, del etilenglicol, para finalmente romper la tensión superficial del agua. Teóricamente, en este punto, la tensión del líquido se romperá y la clasificación del paciente en estudio será de alto riesgo.

Se utilizó la función de transferencia O_k , debido a su comportamiento. Teniendo en cuenta que:

$$\lim_{x \rightarrow \infty} (1 - e^{-x}) = 1 \quad (32)$$

$$\lim_{x \rightarrow 0} (1 - e^{-x}) = 0 \quad (33)$$

Así, cuando la fuerza ejercida por el peso es mayor que la ejercida por el líquido, la tensión superficial se rompe (**Figura 3 - 8**). Cuando esto ocurre, O_k tiende a ser uno, y cuando no lo hace, el valor de O_k tiende a ser cero.

Como tal, los resultados correctos para los cuatro tipos de riesgo deben ser como se muestra a continuación:

$$\left\{ \begin{array}{l} \text{Riesgo 1, } O = \{1, 0, 0, 0\} \\ \text{Riesgo 2, } O = \{1, 1, 0, 0\} \\ \text{Riesgo 3, } O = \{1, 1, 1, 0\} \\ \text{Riesgo 4, } O = \{1, 1, 1, 1\} \end{array} \right. \quad (34)$$

El riesgo 4 rompe todas las tensiones superficiales, mientras que el riesgo 1 solo rompe la primera tensión superficial.

Figura 3 - 8. Clasificación a través de la Red neuronal de tensión superficial.



Fuente: Elaboración propia.

3.6.5. Error de Retropropagación

Las cuatro salidas contenidas en O se comparan con la respuesta E_{ij} , que la red de neuronas debería haber producido, calculando así el error cuadrático medio (35):

$$errorcm = \sum_{k=1}^4 \frac{(O_k - T_k)^2}{2} \tag{35}$$

Los siguientes pasos calculan la influencia de cada parámetro en el error de la red neuronal, a través de la retropropagación del error, a través de derivadas parciales. La siguiente ecuación se deriva de O_k :

3.6.6. La derivada del error, con respecto a la salida de la red neuronal

$$\frac{\partial error_{cm}}{\partial O_k} = \sum_{k=1}^4 (O_k - T_k) \quad (36)$$

3.6.7. La derivada del error, con respecto a la salida de la capa 3

$$\frac{\partial error_{cm}}{\partial Y_1^3} = \frac{\partial error_{cm}}{\partial m} = \frac{\partial error_{cm}}{\partial O_k} * \frac{\partial O_k}{\partial m} \quad (37)$$

$$\frac{\partial O_k}{\partial m} = \frac{g}{T_k * Per} e^{\frac{-m * g}{T_k * Per}} \quad (38)$$

$$\frac{\partial error_{cm}}{\partial Y_2^3} = \frac{\partial error_{cm}}{\partial Per} = \frac{\partial error_{cm}}{\partial O_k} * \frac{\partial O_k}{\partial Per} \quad (39)$$

$$\frac{\partial O_k}{\partial Per} = \frac{-m * g}{T_k * Per^2} e^{\frac{-m * g}{T_k * Per}} \quad (40)$$

$$\frac{\partial error_{cm}}{\partial Y^3} = \left\{ \frac{\partial error_{cm}}{\partial Y_1^3}, \frac{\partial error_{cm}}{\partial Y_2^3} \right\} \quad (41)$$

3.6.8. Derivada del error, según los pesos de la capa 3

$$\frac{\partial error_{cm}}{\partial w_{k,i}^3} = \frac{\partial error_{cm}}{\partial Y_k^3} * \frac{\partial Y_k^3}{\partial w_{k,i}^3} \quad (42)$$

for $k = 1$ to 80, con $i = 1$ y 2

$$\frac{\partial Y_k^3}{\partial w_{k,i}^3} = Y_i^2 \quad (43)$$

3.6.9. Derivada del error, de acuerdo con el sesgo de la capa 3

$$\frac{\partial error_{cm}}{b_k^3} = \frac{\partial error_{cm}}{\partial Y_k^3} * \frac{\partial Y_k^3}{b_k^3} \quad (44)$$

3.6.10. Derivada del error, con respecto a la salida de la capa 2

$$\frac{\partial error_{cm}}{\partial Y_i^2} = \sum_{k=1}^2 \frac{\partial error_{cm}}{\partial Y_k^3} * \frac{\partial Y_k^3}{Y_i^2} \quad (45)$$

for $i = 1$ to 80

$$\frac{\partial error_{cm}}{\partial Y_i^2} = \sum_{k=1}^2 \frac{\partial error_{cm}}{\partial Y_k^3} * w_{k,i}^3 \quad (46)$$

for $i = 1$ to 80

3.6.11. Derivada del error, según los pesos de la capa 2

$$\frac{\partial errorcm}{\partial y^2} = \frac{\partial errorcm}{\partial Y^2} * \frac{\partial Y^2}{\partial y^2} \quad (47)$$

$$\frac{\partial Y^2}{\partial y^2} = 1 - (Y^2)^2 \quad (48)$$

$$\frac{\partial errorcm}{w_{k,i}^2} = \frac{\partial errorcm}{\partial Y^2} * \frac{\partial Y^2}{\partial y^2} * \frac{\partial y^2}{w_{k,i}^2} \quad (49)$$

for $k, i = 1$ to 80

3.6.12. Derivada del error, de acuerdo con el sesgo de la capa 2

$$\frac{\partial errorcm}{b_k^2} = \frac{\partial errorcm}{\partial Y^2} * \frac{\partial Y^2}{\partial y^2} * \frac{\partial y^2}{b_k^2} \quad (50)$$

$$\frac{\partial y^2}{b_k^2} = 1 \quad (51)$$

3.6.13. Derivada del error, con respecto a la salida de la capa 1

$$\frac{\partial errorcm}{\partial Y^1} = \frac{\partial errorcm}{\partial Y^2} * \frac{\partial Y^2}{\partial y^2} * \frac{\partial y^2}{\partial Y^1} \quad (52)$$

3.6.14. Derivada del error, según los pesos de la capa 1

$$\frac{\partial y^2}{\partial Y^1} = w_{k,i}^3 \text{ for } i, k = 1 \text{ to } 80 \quad (53)$$

$$\frac{\partial error_{cm}}{\partial y^1} = \frac{\partial error_{cm}}{\partial Y^2} * \frac{\partial Y^2}{\partial y^2} * \frac{\partial y^2}{\partial Y^1} * \frac{\partial Y^1}{\partial y^1} \quad (54)$$

$$\frac{\partial Y^1}{\partial y^1} = 1 - (Y^1)^2 \quad (55)$$

$$\frac{\partial error_{cm}}{\partial w_{k,i}^1} = \frac{\partial error_{cm}}{\partial Y^2} * \frac{\partial Y^2}{\partial y^2} * \frac{\partial y^2}{\partial Y^1} * \frac{\partial Y^1}{\partial y^1} * \frac{\partial y^1}{\partial w_{k,i}^1} \quad (56)$$

$$\frac{\partial y^1}{\partial w_{k,i}^1} = E_i \text{ for } k = 1 \text{ to } 80; i = 1 \text{ to } 20 \quad (57)$$

3.6.15. Derivada del error, de acuerdo con el sesgo de la capa 1

$$\frac{\partial error_{cm}}{\partial b^1} = \frac{\partial error_{cm}}{\partial Y^2} * \frac{\partial Y^2}{\partial y^2} * \frac{\partial y^2}{\partial Y^1} * \frac{\partial Y^1}{\partial y^1} * \frac{\partial y^1}{\partial b^1} \quad (58)$$

$$\frac{\partial y^1}{\partial b^1} = 1 \quad (59)$$

$$\frac{\partial error_{cm}}{\partial parameters} = \left\{ \begin{array}{l} \frac{\partial error_{cm}}{\partial b^1}, \frac{\partial error_{cm}}{\partial w^1}, \dots \\ \frac{\partial error_{cm}}{\partial b^2}, \frac{\partial error_{cm}}{\partial w^2}, \dots \\ \frac{\partial error_{cm}}{\partial b^3}, \frac{\partial error_{cm}}{\partial w^3} \end{array} \right\} \quad (60)$$

Los nuevos parámetros en la iteración $n + 1$ se calculan a través del método de gradiente conjugado:

$$parametros(n + 1) = parametros(n) + \eta(n) * d(n) \quad (61)$$

Dónde,

$$\eta(n) * d(n) \quad (62)$$

Depende de,

$$\frac{\partial error_{cm}}{\partial parameters} \text{ values.} \quad (63)$$

Este procedimiento se repite, comenzando en el Paso 2.5.1 para los datos de entrenamiento restantes, completando así la primera iteración. Más tarde, las iteraciones se realizan repetidamente hasta que haya una convergencia de la red neuronal artificial, expresada en los siguientes parámetros:

3.6.16. Criterios para la parada de entrenamiento

Para la parada de entrenamiento, se utilizaron los siguientes tres criterios:

- a) Gradiente de rendimiento mínimo, el valor de este gradiente mínimo es 10^{-6} . Esta tolerancia se asigna para el aprendizaje adecuado de la red neuronal;
- b) Rendimiento, para medir el rendimiento de la red neuronal, se emplea el error cuadrático medio. El valor a alcanzar es cero, para evitar presentar errores de salida neural;
- c) Número de iteraciones,

si la red neuronal alcanza las 5.000 iteraciones, detendrá el entrenamiento. Se elige un gran número de iteraciones, ya que, idealmente, se detiene con criterios de error.

El modelo de aprendizaje se evaluó aplicando seis corridas automáticas. Para hacer eso, el conjunto de datos se dividió en grupos de entrenamiento, validación y prueba con diferentes porcentajes de división de datos, como se muestra en la **Tabla 3-5**.

Tabla 3 - 5. División de datos de entrenamiento, validación y testeo.

División de los datos en cada simulación			
Corrida	Entrenamiento	Validación	Test
1	18%	2%	80%
2	36%	4%	60%
3	45%	5%	50%
4	54%	6%	40%
5	72%	8%	20%
6	81%	9%	10%

3.6.17. Medidas estadísticas para la prueba de clasificación

Para la evaluación del algoritmo se utilizaron las siguientes métricas: a) Sensibilidad, que proporciona la probabilidad de que, dada una observación positiva, la red neuronal lo clasifique como positivo (64); b) Especificidad, que proporciona la probabilidad de que, dada una observación negativa, la red neuronal la clasifique como negativa (65); c) Error de clasificación, porcentaje de error contenido en el modelo (66); y, d) exactitud, que proporciona el porcentaje total de exactitud de la red neuronal (67).

$$\text{Sensibilidad} = \frac{TP}{TP + FN} \quad (64)$$

$$\text{Especificidad} = \frac{TN}{TN + FP} \quad (65)$$

$$\text{Error de Clasificación} = \frac{FP + FN}{TP + TN + FP + FN} \quad (66)$$

$$\text{Exactitud} = \frac{TP + TN}{TP + TN + FP + FN} \quad (67)$$

3.7. Paso 3. Análisis de estabilidad

Para analizar la estabilidad del sistema en los resultados obtenidos, se debe realizar un análisis de varianza, utilizando la ecuación (68), para establecer si existen diferencias significativas en los resultados.

En este análisis, representando la respuesta a las variables, T_i , fue el efecto causado por el enésimo tratamiento, y ε_i , el enésimo error experimental. La información recopilada debe cumplir con los requisitos de independencia y normalidad. El análisis de varianza se realizó bajo un intervalo de confianza del 99,5% (Rodríguez, 2007):

$$Y_i = \mu + T_i + \varepsilon_i \quad (68)$$

El código desarrollado en la versión de Matlab se puede consultar aquí y al final en el anexo I:

<https://codeocean.com/capsule/6532855/tree/v1> (Mosquera et al., 2019a). La versión 9.4 de Matlab fue usada para el desarrollo del algoritmo.

3.8. Paso 4. Clasificación del nivel de riesgo psicosocial y recomendaciones para su prevención

Después de procesar la predicción de los niveles de riesgo psicosocial, el algoritmo genera una serie de indicaciones para actuar sobre el individuo que presente niveles altos y muy altos de riesgo siguiendo lo indicado en la literatura y en las normas de Seguridad y Salud en el Trabajo según la (Guía Técnica General del Ministerio del Trabajo; Promoción, Prevención e Intervención de los Factores Psicosociales y sus Efectos en la Población Trabajadora, 2015).

3.9. Comparación con otras técnicas inteligentes

Los siguientes pasos muestran pruebas previas con otros modelos para evaluar los resultados de clasificación con respecto al nivel de riesgo psicosocial.

3.9.1. Comparación con el algoritmo j48 árbol de decisión y naive bayes

Mosquera et al., (2016b) realizaron pruebas con éstos algoritmos siguiendo los siguientes pasos para evaluar su eficiencia en la calidad de la predicción respecto al nivel de riesgo psicosocial en docentes de colegios públicos:

Paso 1: Limpieza de los datos y pre-procesamiento de la información.

Paso 2: Reducción de las variables consideradas datos atípicos a través de algoritmos genéticos.

Paso 3: Filtrado de atributos, para refinar los resultados de los modelos.

En primera instancia los datos fueron evaluados a través del algoritmo Naive Bayes:

Donde la búsqueda diagnóstica más probable, C^* , está dada por (X_1, \dots, X_n) los síntomas para un paciente determinado, donde \mathbf{p} , es un factor psicosocial asociado que depende solo de X_1, \dots, X_n , es decir, los valores de X_i deben ser conocidos (Y deben ir desde 0 hasta 4, el grado de riesgo psicosocial).

$$= \mathit{arg}_{c} \mathit{máx} p(C = c) \prod_{i=1}^n p(X_i = x_i / C = c) \quad (69)$$

Paso 4: Análisis de los datos a través del algoritmo árboles de decisión j48.

Los datos de entrenamiento son un sistema $S = s_1, s_2, \dots$ de muestras de variables psicosociales clasificadas. Cada muestra $S_i = \{x_1, x_2, \dots\}$ es un vector, en el que x_1, x_2, \dots representa los atributos o características del riesgo psicosocial. Los datos de entrenamiento aumentan con el vector $C = \{c_1, c_2, \dots\}$ donde c_1, c_2, \dots representan la clase a la que pertenece cada individuo expuesto al riesgo psicosocial. El mejor rendimiento de predicción, obtuvo un 91 y 89%, respectivamente.

3.9.2. Comparación con la red neuronal artificial de retro-propagación Perceptrón utilizando algoritmo genético para reducir la dimensionalidad de los datos

Mosquera et al., (2018b) utilizaron algoritmos genéticos y análisis de componentes principales para reducir la dimensionalidad de los datos de un conjunto de datos de evaluaciones psicosociales para comparar los resultados utilizando redes neuronales perceptron de retropropagación siguiendo los siguientes pasos:

La capa de entrada, con 123 variables de entrada, las capas ocultas (compuestas por 12 neuronas) y cuatro capas de salida (riesgo muy alto, alto, medio y bajo). En segundo lugar, en la capa de entrada, el número de variables a ingresar fue reemplazado por solo 20 variables, haciendo que las capas ocultas y las capas de salida sean iguales a las iniciales. Cada una de las columnas de datos representa los valores $R_{i,1} \dots R_{i,m}$, que se ingresan a cada nodo. Posteriormente, un $W_{i,1} \dots W_{i,m}$, como función de peso definida. La ecuación para modelar el riesgo psicosocial está dada:

$$f_i(x) = \left(\sum_{i=1}^n W_{i,j} * FR_{i,j} \right) \quad (70)$$

Donde, para cada uno de los nodos de red $N_1 \dots N_n$, se define la función de transferencia que se muestra en la Ecuación (65). Allí, n es el número de individuos evaluados, $FR_{i,j}$ es el valor del riesgo i de cada individuo con respecto a cada variable j , y $W_{i,j}$ es el peso calculado para los índices de riesgo i de cada individuo j .

Para hallar la mejor población de datos que alimente la red neuronal, se definió una función fitness para el algoritmo genético (71). En este caso, esta función entrena el clasificador, utilizando las 20 características contenidas en los padres, con la información de las muestras extraídas de la base de datos. El mejor rendimiento de predicción obtuvo un 93.%.

$$Fitness_{G.Risk} = \min \left\{ \sum_{i=1}^N P_{ij} \right\} \quad (71)$$

Donde, " i " representa cada variable, " N " representa el número total de variables con sus respectivos valores, y " P_{ij} " se refiere a la eficiencia de cada variable i con respecto al grado de riesgo j . Las subpoblaciones corresponden a un grupo de 10 individuos, donde cada individuo representa un grupo de 20 características.

"Individuo $_i$ = {grupo de 20 variables de 123}, ($i=1,2,3...,10$)".

3.9.3. Comparación con máquinas vectoriales (svm lineal)

Mosquera et al., (2018a) realizaron un estudio comparativo con el modelo de aprendizaje automático para la predicción: Máquina de soporte vectorial (MSV). Se seleccionaron las 20 mejores variables en este estudio (X_{ij}), con la mayor eficiencia de exactitud de predicción.

El mejor rendimiento de predicción obtuvo un 93.3%. Para calcular las clases de predicción, se utilizó el siguiente clasificador multiclase svm (Schlkopf et al., 2018):

$$\underset{w_r \in K, \epsilon^r \in R^n, b_r \in R}{\text{minimize}} \frac{1}{2} \sum_{r=1}^N \|w_r\|^2 + \frac{C}{n} \sum_{i=1}^n \sum_{r \neq y_1} \epsilon_i^r, \quad (72)$$

Dónde, $n \in \{1, \dots, N\} \setminus Y_i$ y $Y_i \in \{1, \dots, N\}$ era la etiqueta del patrón multiclase X_{iRisk} del atributo riesgo perteneciente a cada individuo.

3.9.4. Comparación con máquinas de soporte vectorial Rbf y algoritmo híbrido ascenso de colina-máquinas de soporte vectorial (hc-svm)

Mosquera Navarro et al., (2018) utilizaron el algoritmo híbrido de ascenso de colina-máquinas de soporte vectorial (HC-SVM). Aquí, el proceso de usar la heurística llamada ascenso de colina, permite encontrar el óptimo local dentro del espacio de búsqueda de los valores X_{ij} (atributos de los factores de riesgo psicosocial), con el fin de obtener los parámetros de clasificación óptimos para el algoritmo de Soporte Vectorial en el menor tiempo posible de convergencia. El algoritmo de ascenso de colina es un algoritmo de búsqueda local, que realiza iteraciones en la vecindad local de la forma $R_n(m(i), m) = 0$ para todo $m(i) \in \Omega$, $m \in \eta(m(i))$ que detiene las iteraciones cuando no encuentra opciones de iteración más óptimas (Jacobson & Yücesan, 2004). El uso del algoritmo híbrido de escalada y la máquina de vectores de soporte (HC-SVM) proporcionó una clasificación adecuada del 95% del nivel de riesgo psicosocial.

3.9.5. Comparación con el algoritmo híbrido knn-máquina de vectores de soporte (knn-svm)

Mosquera et al., (2019) utilizaron el algoritmo híbrido de k vecinos próximos-máquinas de soporte vectorial propuesto por Zhou et al., (2013) para ser adaptado para evaluar el desempeño de clasificación del algoritmo con respecto al grado riesgo psicosocial en docentes de colegios públicos obteniendo un porcentaje de clasificación del 86.66%.

V. Capítulo 4 - Resultados y discusión – Comprobación de la hipótesis de la investigación

4.2. Desarrollo de algoritmos previos a la obtención del modelo propuesto

En esta etapa del proceso antes de llegar al desarrollo del algoritmo propuesto se evaluaron los resultados de dos algoritmos previos desarrollados y evaluados con parte del conjunto de datos (495), para observar la adaptabilidad al tipo de datos que se manejaron y la respuesta de cada algoritmo para ir construyendo el camino que finalmente nos condujo al algoritmo de red neuronal de tensión superficial.

4.2.1. Algoritmo de ascenso de colinas con reinicio aleatorio – Máquinas de Soporte Vectorial (RRHC-SVM)

Éste modelo utilizó las ventajas de optimización y búsqueda heurística de reinicio aleatorio para hallar el mejor óptimo local que permitiera al algoritmo de clasificación mejorar en cada reinicio la clasificación vectorial. Los resultados obtenidos a través de la matriz de confusión y de exactitud promedio para éste modelo fueron muy similares a los de una red

neuronal con características de retropropagación. Los valores de clasificación promedio en la matriz de confusión para cada vector fue de 92% y de exactitud promedio de 93,4%. El algoritmo desarrollado se puede observar en el anexo G de éste documento.

4.2.2. Algoritmo de búsqueda cucko – Máquinas de Soporte Vectorial (CS-SVM)

Ésta metaheurística bioinspirada en los nidos de cucko y en el vuelo de levy para realizar la búsqueda del óptimo global desarrollada por Yang & Suash Deb, (2009) utilizó las ventajas de hallar el óptimo global en el espacio de búsqueda para entregar los mejores individuos para lograr un adecuada clasificación. Los resultados obtenidos a través de la matriz de confusión y de exactitud promedio para éste modelo de clasificación promedio en la matriz de confusión fue de 92,4% y de exactitud promedio de 93%. El algoritmo desarrollado se puede observar en el anexo H de éste documento.

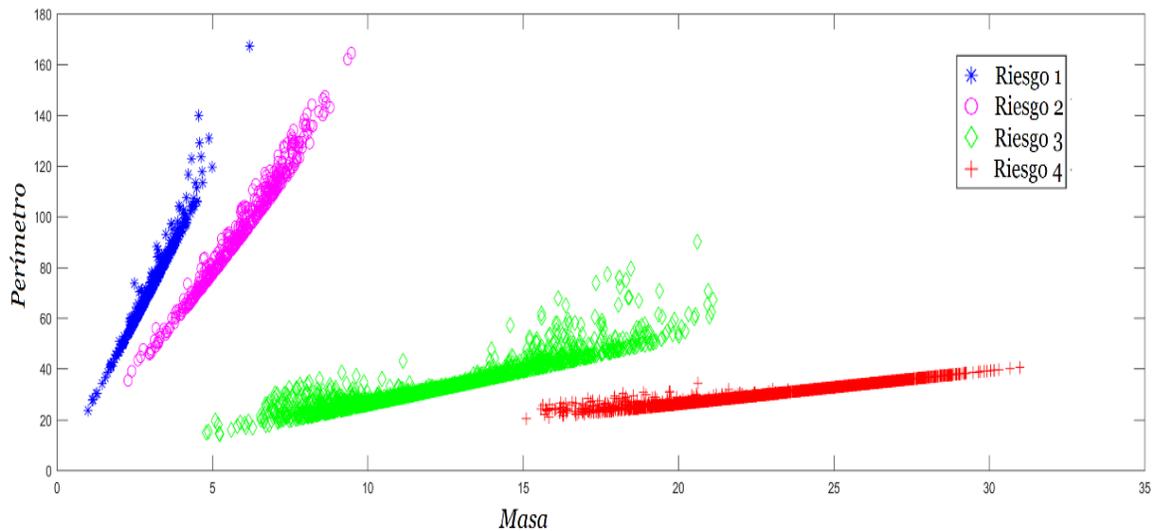
4.3. Paso 1 y 2. Algoritmo de red neuronal de tensión superficial física

En éste capítulo se muestran los resultados obtenidos del modelo desarrollado basado en la teoría de la tensión superficial física (Macleod, 1923) adaptado a una red neuronal de retropropagación de cuatro capas en total.

4.3.1. Masa(m) y Perímetro(Per)

La **Figura 4 - 1** muestra los resultados detectados en la capa de salida que representa el modelo de tensión superficial (**Figura 3 - 6**), aplicado a las 20 variables de los 5443 pacientes objeto de estudio. La figura representa los resultados del perímetro y las salidas de masa para cada sujeto. Hay un nivel de riesgo separado en áreas específicas, y los tipos de riesgo pueden interpretarse adicionalmente en forma física. De hecho, el mayor riesgo en el gráfico corresponde a las cruces rojas, que presentan valores de masa que, en términos relativos, son más grandes que el resto, junto con perímetros relativamente más pequeños, que provocan la ruptura de la tensión superficial de los cuatro líquidos. El riesgo más bajo (representado en azul con asteriscos) presenta perímetros relativamente altos y masas relativamente bajas, lo que hace que permanezca en la superficie.

Figura 4 - 1. Modelo de clasificación del riesgo perímetro vs. Masa.

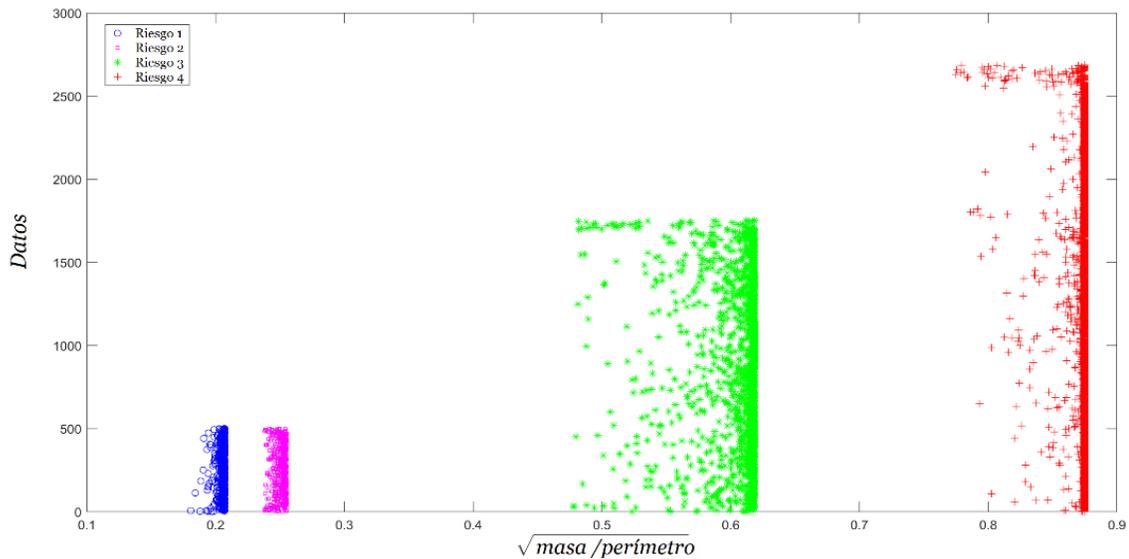


Fuente: Elaboración propia.

4.3.2. Rendimiento del modelo de red neuronal de tensión física superficial

En la **Figura 4 - 2**, se representa la raíz cuadrada de la relación masa / perímetro. La raíz se aplica solo para mejorar la visualización de las separaciones entre los riesgos. Éste nivel de riesgo varía entre los extremos de esta relación. Si se grafica la relación masa-perímetro, se pueden observar los datos mejor separados, como se muestra a continuación:

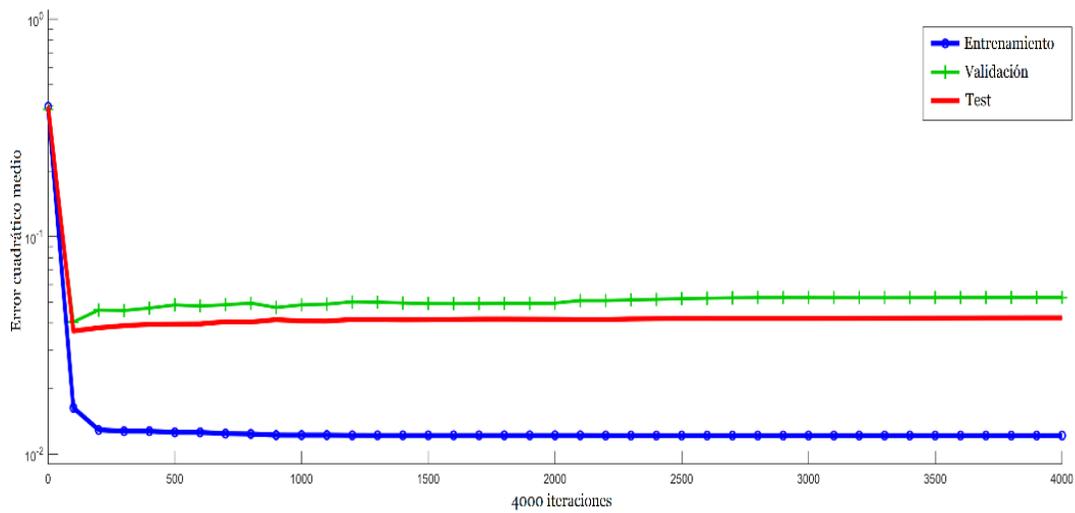
Figura 4 - 2. Visualización de la separación del nivel de riesgo en el modelo.



Fuente: Elaboración propia.

La **Figura 4 - 3** muestra el rendimiento de la red neuronal de tensión superficial, con el paso de las iteraciones. La línea de entrenamiento azul disminuye, ya que el objetivo de la red neuronal de tensión superficial es ajustar sus parámetros para facilitar la disminución del error de iteración. Por esta razón, en los primeros 50, las curvas disminuyen juntas.

Figura 4 - 3. Rendimiento de iteraciones en el modelo de red neuronal de tensión superficial.



Fuente: Elaboración propia.

La **Tabla 4 - 1** muestra el grupo de datos consolidados que se evaluó con cada una de las métricas descritas, con seis ejecuciones diferentes, cada una con diferentes valores de entrenamiento, validación y prueba.

Tabla 4 - 1. Resultados para el conjunto de datos.

Simulación	Sensibilidad	Especificidad	Error de Clasificación	Exactitud
1 (18-2-80)	83.31%	96.28%	5.13%	94.87%
2 (36-4-60)	85.23%	96.60%	4.60%	95.40%
3 (45-5-50)	85.23%	96.60%	4.60%	95.40%
4 (54-6-40)	89.21%	97.46%	3.35%	96.65%
5 (72-8-20)	90.92%	97.97%	2.63%	97.37%
6 (81-9-10)	89.24%	97.59%	3.18%	96.82%

Los resultados muestran que, para la primera simulación, con un 18% de entrenamiento, un 2% de validación y un 80% de datos de prueba, la sensibilidad del algoritmo arroja una sensibilidad del 83.31% y 94.87% de exactitud, mientras que, para un sexto grupo de datos, con un 72% de entrenamiento, 8 % de validación y 20% de datos de prueba, la sensibilidad arroja un resultado de 90.92% y 97.37% de exactitud, comportándose como la mejor opción para la clasificación de resultados.

4.3.3. Matriz de confusión para el modelo de Red neuronal de tensión superficial

Esta matriz de confusión específica (**Tabla 4 - 2**) muestra la visualización del rendimiento del algoritmo RNA-TS.

Tabla 4 - 2. Matriz de confusión del modelo de Red neuronal de tensión superficial para la predicción del riesgo psicosocial, para el conjunto de datos de entrenamiento (72%), validación (8%) y prueba (20%).

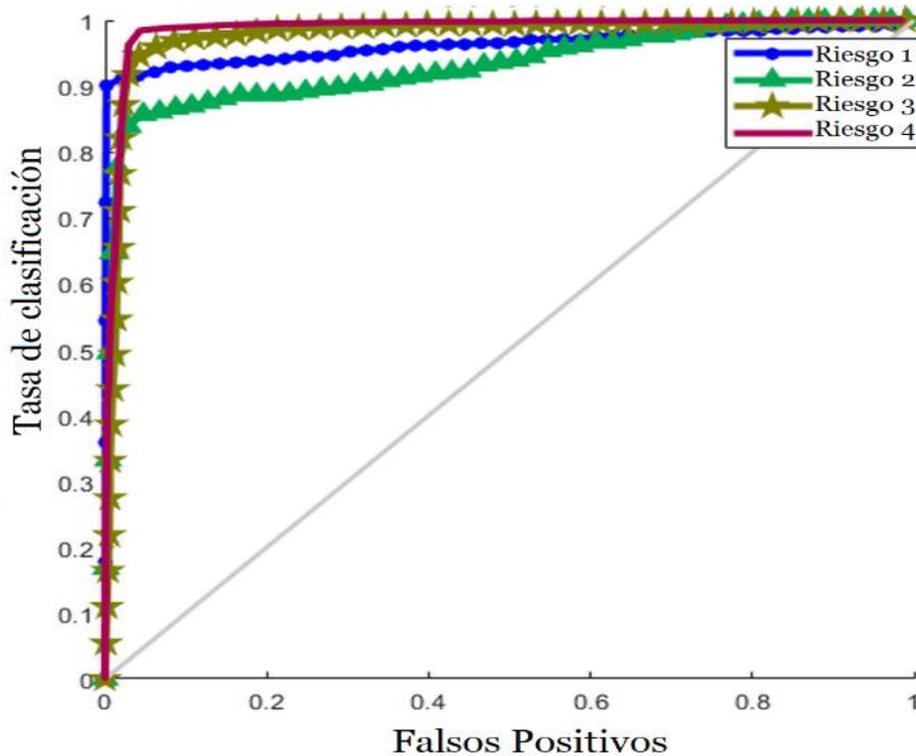
		Riesgo 1	Riesgo 2	Riesgo 3	Riesgo 4		
Salida Clase	Riesgo 1	492	5	4	0	98.2%	
		9.0%	0.1%	0.1%	0%	1.8%	
	Riesgo 2	16	439	44	1	87.8%	
		0.3%	8.1%	0.8%	0.0%	12.2%	
	Riesgo 3	24	53	1643	35	93.6%	
		0.4%	1.0%	30.2%	0.6%	6.4%	
	Riesgo 4	14	50	44	2579	96.0%	
		0.3%	0.9%	0.8%	47.4%	4.0%	
			90.1%	80.3%	94.7%	98.6%	94.7%
			9.9%	19.7%	5.3%	1.4%	5.3%
			Riesgo 1	Riesgo 2	Riesgo 3	Riesgo 4	
			Clase objetivo				

La exactitud del resultado promedio más alto y más bajo fue 98.6% (Riesgo 4) y 80.3% (Riesgo 2) respectivamente, como resultado de la división de datos de entrenamiento, validación y prueba.

4.3.4. Curva ROC

Como se muestra en la curva ROC (**Figura 4 - 4**), el riesgo 4 presenta la mejor clasificación de todos los riesgos (98.6%), que coincide con la matriz de confusión, mientras que el riesgo 2 es el que presenta la mayor confusión en la clasificación (19.7%).

Figura 4 - 4. Curva ROC para la clasificación del riesgo psicosocial del modelo Red neuronal de tensión superficial.



Fuente: Elaboración propia.

De acuerdo con esta curva ROC, el modelo de Red neuronal de tensión superficial puede detectar el 98.6% de los casos con nivel de riesgo psicosocial Muy Alto mientras genera una tasa de falsos positivos de 1.4%.

La **Tabla 4 - 3** muestra que el nivel de riesgo 2 tuvo la sensibilidad más baja (80.26%), lo que indica que el nivel de riesgo 2 fue el tipo de riesgo más difícil de clasificar. De hecho, un sujeto perteneciente al nivel de riesgo 2 se clasificaría como tal por la red neuronal de tensión superficial con una probabilidad del 80,26%. Además, el nivel de riesgo 2 tenía la exactitud más baja, lo que indica que la red neuronal de tensión superficial clasificaría correctamente a un individuo con una probabilidad del 96,96%.

Por el contrario, para los individuos que presentaron el nivel de riesgo 4, la red neuronal de tensión superficial presentó una probabilidad de clasificación del 97.37% en dicho tipo de riesgo y un valor de sensibilidad del 98.62%. Sin embargo, el nivel de riesgo con mayor exactitud fue el Nivel de riesgo 1. Cuando la red neuronal de tensión superficial supuso que un individuo pertenecía al tipo de riesgo uno, la probabilidad de que esto fuera cierto fue del 98.85%.

Tabla 4 - 3. Medidas estadísticas para la prueba de clasificación para los cuatro niveles de riesgo. Para el conjunto de datos de entrenamiento (72%), validación (8%) y prueba (20%).

Medidas Estadísticas	Niveles de Riesgo			
	Nivel Riesgo 1	Nivel Riesgo 2	Nivel Riesgo 3	Nivel Riesgo 4
Sensibilidad	90.11%	80.26%	94.70%	98.62%
Especificidad	99.82%	98.78%	97.05%	96.23%
Exactitud	98.85%	96.96%	96.31%	97.37%
Error Clasificación	1.15%	3.04%	3.69%	2.63%

4.4. Paso 3. Análisis de estabilidad

Consiste en permitir la evaluación del algoritmo hasta que no sea posible encontrar una solución mejor. El primer proceso se repitió 7 veces (tratamientos), tomando en cada tratamiento los diez mejores resultados (**Tabla 4 - 4**).

Tabla 4 - 4. Porcentaje de exactitud para resultados predictivos sobre el nivel de riesgo psicosocial en maestros de escuela colombianos.

Replicas	DT-J48	NB	RNA	MSV	MSV RBF	HC-SVM	RNA-TS
1	89	89	92	92	89.26	92	96.97
2	91	89	92.5	93	89.27	93	97.37
3	91	90	93	93	89.26	93	97.37
4	92	90	93	93	89.26	93	97.37
5	92	90	93	93	89.27	93	97.37
6	92	90	93	93	89.26	93	97.37
7	92	90	93	93	89.27	93	97.37
Promedio	91.29	89.71	92.83	92.86	89.26	92.86	97.31

Nota: DT: Árboles de decisión J48; NB: Naïve Bayes; RNA: Red Neuronal Artificial; MSV: Máquinas de Soporte Vectorial Lineal; MSV-RBF: Máquina de Soporte Vectorial de Base Radial; HC-SVM: Ascenso de Colina-Máquina de Soporte Vectorial y RNA-TS: Red neuronal de tensión superficial.

La **Tabla 4 - 5** contiene el análisis de varianza con un intervalo de confianza del 99.5%, usando las soluciones encontradas en la **Tabla 4 - 4**. Dado que, en la **Tabla 4 - 5**, el F tabulado es mayor que el F calculado, se puede decir que hay diferencias significativas en los resultados dados por el sistema.

Tabla 4 - 5. Análisis de varianza (Anova).

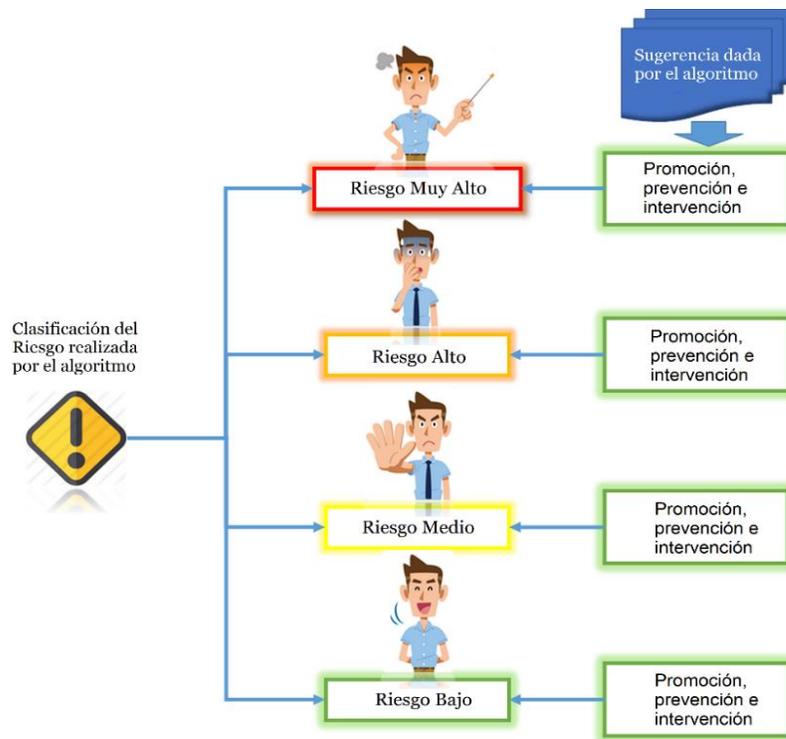
Variación	SS	df	MS	F	Prob>F
Columns	0.0969	9	0.01077	0.62	0.7778
Error	1.5650	90	0.01739		
Total	1.6619	99			

Fuente: Elaboración propia.

4.5. Paso 4. Clasificación del nivel de riesgo psicosocial y recomendaciones para su prevención

La obtención de la clasificación del nivel de riesgo en la que se clasifica al individuo evaluado, a través del algoritmo híbrido, genera ciertas recomendaciones, basadas en los casos previamente documentados y que están contenidos en la literatura científica, que describe el manejo del individuo, basado en el nivel percibido (**Figura 4 - 5**). Se desarrolló una interfaz para facilitar al usuario que evalúa a las personas interactuar con el algoritmo, generando indicaciones basadas en la guía técnica de intervención psicosocial (Ministerio del trabajo, 2015) (Anexo F).

Figura 4 - 5. Flujo de recomendación para prevención del riesgo psicosocial basado en el algoritmo Red neuronal de tensión superficial.



Fuente: Elaboración propia.

4.6. Resultados experimentales basados en métodos de comparaciones previas

En esta sección, los autores comparan y muestran los resultados obtenidos previamente con la aplicación de técnicas de clasificación (**Tabla 4 - 6**). Finalmente, las características psicosociales predictivas, con los algoritmos muestran diferente efectividad (Mosquera et al., 2018; Mosquera, Castrillón, et al., 2018c; Mosquera et al., 2016, 2018, 2019).

Tabla 4 - 6. Resultados aplicando diferentes técnicas de clasificación. *Exactitud.

ítem	Algoritmo	% promedio clasificación*
1	J48	91
2	Naïve Bayes	89
3	ANN	93
4	Decision Tree	90
5	SVM	93
6	RRHC-SVM	93.84
7	CS-SVM	94.90
8	SVM-RBF	89.26
9	KNN	84.66
10	KNN-SVM	86.66
10	Método propuesto: RNA-TS	97.31

Fuente: Elaboración propia.

4.7. Discusión de resultados

4.7.1. Aspectos generales

Una adecuada metodología de identificación de riesgo psicosocial con técnicas inteligentes, debe contemplar un número adecuado de parámetros (que para este caso fue de 20 por persona) el cual permita la adaptabilidad de la metodología a otras regiones del país teniendo en cuenta los factores sociales, económicos, ambientales, culturales, educativos, que permiten mejorar el grado de predicción del algoritmo.

En esta investigación el valor total de parámetros adecuado para el diseño y adaptabilidad del algoritmo para garantizar su mejor rendimiento fue de 20, con significancia estadística importante, la identificación de éste número de variables se logró gracias a la reducción de la dimensionalidad de los datos a través de la aplicación del algoritmo ReliefF (Robnik-S̃ & Kononenko, 2013).

Igualmente, en la literatura científica no se reconocen estudios de aplicación de técnicas de inteligencia artificial únicamente en el constructo teórico de riesgo psicosocial aquí estudiado, la mayoría de estudios identificados realizan aproximaciones a los factores de riesgo psicosocial asociado a otras teorías como lo son, la satisfacción laboral, el estrés, la salud física, la salud mental, la accidentalidad laboral, el ausentismo laboral debido a factores psicosociales, la deserción laboral asociada a factores psicosociales, la rotación de empleados asociada a los factores psicosociales, por lo tanto, la mayoría de resultados donde se introducen técnicas inteligentes siempre llevaran como variable asociada el factor psicosocial.

El problema de identificación con respecto a los riesgos psicosociales impuesta en estos días en los ambientes laborales, es de interés relevante debido a que se orienta a

consolidar la eficiencia y la competitividad de las organizaciones educativas públicas basado en el bienestar laboral de los docentes.

La metodología propuesta en ésta investigación se aplicó a la predicción del riesgo psicosocial Intralaboral. La selección final de los parámetros que se tuvieron en cuenta para el desarrollo del algoritmo se obtuvieron de pruebas realizadas a lo largo de la investigación como, por ejemplo, análisis de correlación, análisis de PCA, algoritmos genéticos, algoritmo de la cantidad esperada de la información, algoritmo GainRatioAttributeEval.

En esta investigación se referenciaron 20 parámetros, los cuales pertenecen a características psicosociales, fisiológicas y físicas por persona, y se estableció una prueba de anova con una confiabilidad del 95%, con el fin de determinar la confiabilidad de los resultados.

Para garantizar los resultados adecuados estadísticamente del algoritmo propuesto, también se realizaron validaciones cruzadas con el conjunto de datos (cross-validation).

Con relación al estado del arte en riesgos psicosociales (British Standards Institution (BSI), 2011; Canadian Standards Association, 2013; European Agency for Safety and Health at Work, 2012; Stavroula Leka et al., 2015; Memish et al., 2017), los procesos de investigación de los factores de riesgo psicosocial utilizan técnicas fundamentadas en la sociología, la psicología laboral, la organización y administración, la organización del trabajo y la salud laboral, no se identifica de manera fehaciente a nivel mundial que se utilicen técnicas de inteligencia artificial para la identificación y clasificación de riesgo psicosocial en los sistemas de gestión de la Seguridad y Salud en el Trabajo.

El número de variables con significancia estadística importante (20) guarda cierta similitud en algunas variables con lo que plantea la guía para la gestión de riesgos psicosociales de la unión europea PAS1010 (Leka & Cox, 2011) que define 13 variables para la evaluación

de los riesgos psicosociales a nivel europeo (Contenido de trabajo, violencia en el trabajo, acoso laboral, desarrollo profesional, condiciones ambientales y riesgos en maquinaria y equipos, cultural organizacional y funciones, rol en el cargo, relaciones interpersonales en el trabajo, horarios de trabajo, control sobre el trabajo, interface hogar-trabajo y carga y ritmo de trabajo).

Actualmente, el modelo de evaluación de riesgos psicosociales europeo (Leka et al., 2017, p. 3) tiene en cuenta 10 dimensiones para evaluar los aspectos negativos psicosociales relacionados con el trabajo (funciones y cultura organizacional, contenido de trabajo, carga y ritmo de trabajo, horarios de trabajo, control, ambiente de trabajo y equipos, relaciones interpersonales en el trabajo, desarrollo profesional e interacción trabajo-hogar), acorde a éste modelo, los hallazgos relativos a la evaluación psicosocial del trabajo en el sector educación por parte de la significancia estadística de nuestra investigación toma relevancia, ya que las variables utilizadas se obtuvieron a través de la aplicación del algoritmo ReliefF y de la reducción de la dimensionalidad de los datos, mostrando su importancia y guardando semejanzas y diferencias entre los contextos europeo y colombiano.

Los resultados de las veinte (20) variables propuestas aquí pueden ser mejorados logrando una mayor participación de los docentes de la región evaluada, logrando tener un volumen mayor de datos diferentes a los psicosociales, similares a los del trabajo realizado por (Suárez Sánchez et al., 2011) que logra hallar resultados relevantes en las correlaciones entre los factores psicosociales y la accidentalidad laboral a través del modelo de máquinas de soporte vectorial, teniendo como base el conjunto de datos de la sexta evaluación nacional de condiciones del trabajo en España, esto puede ser debido a las condiciones y características de índole social, económico, ambiental, organizacional, macroergonómico y cultural de un país de primer mundo.

Las particularidades de cada región colombiana tienden a afectar de manera significativa los resultados de predicción, las variables de riesgo psicosocial pueden variar de región a región incidiendo en los resultados y en el rendimiento del algoritmo.

La aplicación de la metodología en otras regiones del país es fácil de implementar, se requiere obtener el conjunto de datos con el cual entrenar el algoritmo. También, la metodología se puede aplicar a otros sectores y subsectores de la economía teniendo en cuenta las características de cada sector, lo que ayudaría en gran manera a tener un conjunto de datos significativo para realizar predicción con el algoritmo en las diferentes empresas de las regiones y ayudando a los profesionales a mejorar las estrategias de prevención.

El desarrollo y evaluación de éste algoritmo se realizó en la versión de software Matlab V9.8. Éste código puede ser consultado en: <https://codeocean.com/explore/capsules?query=rodolfo%20mosquera> (Mosquera et al., 2019a)

4.7.2. Aspectos específicos

Se ha desarrollado un modelo de predicción y clasificación basado en la teoría de la tensión superficial denominado red neuronal de tensión superficial (RNA-TS), y se han realizado comparaciones con los siguientes modelos: modelos de árbol de decisión (algoritmo J48), Naïve Bayes (NBC), redes neuronales artificiales (ANN), máquinas de vectores de soporte (SVM), Modelos de función de base radial de máquina de vectores de soporte (SVM-RBF), Máquinas de vectores de soporte de escalada (HC-SVM), Máquinas de vectores de soporte de búsqueda de cuco (CS-SVM) y el Vector de soporte vecino más cercano k - modelo híbrido de máquina de soporte (k-NN-SVM).

Para el reconocimiento de patrones y el nivel de predicción de riesgo psicosocial, el modelo de red neuronal de tensión superficial se presenta con bondades frente a la metodología de recolección de datos individual (MS Excel), debido a sus robustas habilidades explicativas y selección de atributos de alta dimensionalidad, a la economía de tiempo en la evaluación y en la obtención de los resultados para identificar y clasificar el riesgo en un

trabajador. Sus aplicaciones, desarrolladas para la administración gerencial y el reconocimiento de patrones en entornos de trabajo y climas organizacionales, son vitales debido al impacto que puede generar en la prevención del riesgo en las organizaciones.

El punto fundamental de la estructura de este modelo es la adaptación de la tensión superficial física a la mejora del modelo de red neuronal, logrando beneficiarse de las características de las redes neuronales y aumentar la exactitud a través de la innovación, en forma de modificación de la red neuronal. Dicha combinación permite la eliminación de deficiencias del modelo lineal y el desarrollo de un enfoque con aspectos del mundo real. Esta aplicación fue probada mediante la validación del modelo en docentes de secundaria, lo que permitió la predicción rápida del riesgo, los coordinadores evaluaron parcialmente el modelo desarrollado y los psicólogos laborales que realizaron evaluaciones psicosociales del docente para resaltar la utilidad y el beneficio del mismo, en términos de la disminución en el tiempo necesario para clasificar los tipos de riesgos que presenta cada docente. Estas dos opiniones son vitales para establecer la importancia del desarrollo de este tipo de herramientas tanto en el ámbito de la salud ocupacional nacional como a nivel mundial, y destaca el liderazgo de Colombia en Iberoamérica en el desarrollo de herramientas que contribuyen al campo de la salud y la seguridad en el trabajo.

En los experimentos llevados a cabo en ésta investigación, los porcentajes de entrenamiento, validación y evaluación fueron 72%, 8% y 20%, respectivamente (**Tabla 4 - 1**). Con estos, se lograron mejores resultados, en comparación con las otras configuraciones evaluadas, así como con los otros valores dados para los mismos procesos. La evaluación sistemática de la validación cruzada reveló los mejores resultados, como se muestra en la **Tabla 4 - 3**, con respecto a las medidas de evaluación del modelo.

Los resultados de la matriz de confusión (**Tabla 4 - 2**), demuestran que el modelo desarrollado es altamente eficiente, en términos de clasificación de riesgo psicosocial, en comparación con otros experimentos similares como en el caso de (Aliabadi, 2015; Larrabee et al., 2003) quienes utilizaron regresión logística y regresión múltiple paso a

paso para predecir la satisfacción laboral de enfermeras de un hospital a través de la correlación de variables, al igual que (Baradaran et al., 2008) quienes utilizaron PCA y el modelo multiobjetivo para identificar la satisfacción en el trabajo y hallar como plantear soluciones para incrementar las condiciones laborales asociadas a factores psicosociales. (Yigit & Shourabizadeh, 2017) quienes utilizaron técnicas como regresión logística, arboles de decisión y máquinas de soporte vectorial para hallar la predicción de rotación de empleados asociada a la influencia de factores psicosociales obteniendo resultados de 88% y 87% de clasificación para algoritmos como máquinas de soporte vectorial y de bosque aleatorio respectivamente.

Los valores obtenidos en la **Tabla 4 - 5**, demuestran la evaluación estadística del modelo de predicción y confirman, mediante la técnica de Anova, que los resultados son estadísticamente estables y confiables. Por lo tanto, en este caso, el modelo de predicción es adecuado.

Posada-Quintero et al., (2020) encontraron en los resultados de la investigación realizada con datos de riesgo psicosocial de docentes Colombianos, que los árboles de decisión logró identificar el 92% de los factores de riesgo y el 71% de los síntomas, con máquinas de soporte vectorial logró identificar el 96% de los riesgos y el 67% de los síntomas, con el algoritmo RUSBoostDT logró identificar el 80% de los riesgos, pero el 82% de los síntomas.

Los síntomas asociados en éste modelo de aprendizaje de máquinas fueron la fatiga, el dolor de cabeza, la irritabilidad, las ganas de llorar, la pérdida del apetito, el sueño irregular, la depresión, la dificultad en la comunicación, el agotamiento mental, los dolores musculoesqueléticos, problemas de voz y problemas gástricos. Comparado con nuestro modelo el cual logro reconocer el grado de riesgo con un 97.31%, se muestra una ventaja significativa en el reconocimiento del grado de riesgo psicosocial y los resultados de los síntomas que se utilizaron muy similar a los utilizados en éste estudio, con hallazgos significativos para lograr hacer énfasis de prevención en las variables que afectan a los docentes.

Los experimentos realizados por González Fuentes et al., (2020) asociados a factores que ocasionaron ausentismo basados en variables ergonómicas, psicosociales y del trabajo, aplicando Regresión Adaptativa Multivariada y tomando los datos de la sexta encuesta europea de las condiciones del trabajo arrojaron resultados del valor medio del R² para los 1000 modelos entrenados para las variables finalmente seleccionadas en cada caso fue de 74,26% con un error cuadrático medio de 27,51, el modelo con los mejores resultados obtenidos fue el de máquinas de soporte vectorial con valores de 67.32% de valor medio y con un error cuadrático medio de 29.01. Datos bastante alejados de los obtenidos en nuestro modelo, y en cuanto a sintomatología identificada, las tareas repetitivas, el dolor de cabeza, la reestructuración organizacional, el trabajo por fuera de los horarios de trabajo y la insatisfacción con las condiciones laborales fueron los principales hallazgos. Se puede ver que se presenta cierta relación en cuanto a las variables manejadas en nuestra investigación.

Los hallazgos del trabajo realizado por Uronen et al., (2020) en enfermeras, mostró que la carga de trabajo y el ritmo de trabajo son factores psicosociales que a través del modelo de redes neuronales pudieron identificar con relativa efectividad. Comparado con nuestro modelo el cual tiene como variable la demanda de carga mental, puede aproximarse a los resultados obtenidos por ellos en esa variable, pero con mucho mejor procesamiento de las variables y aproximación del modelo matemático nuestro método logra obtener mejores resultados.

El estudio desarrollado por Márquez Gómez, (2020) asocio las mayoría de problemas musculoesqueléticos que pudo observar a problemas de control por parte de los trabajadores, dolores en hombros, espalda general, manos, muñecas, cuello, piernas, pies, codos y antebrazos fueron seleccionados como variables predictoras, los modelos utilizados para lograr la predicción fueron árboles de decisión, regresión logística y un modelo de árboles logístico, el modelo con mejor aproximación en la predicción fue el de árboles de decisión con un 71%. Comparado con nuestra metodología, se tiene que converge en la aproximación y asociación que se hace de los problemas musculoesqueléticos y como ellos se convierten en un fenómeno inicial para poder

identificar el riesgo psicosocial, casi todas las variables utilizadas en esta investigación se usaron en la investigación de Márquez.

Jebelli et al., (2019) utilizaron una red neuronal convolucional profunda de 83 neuronas en la primera capa oculta y de 23 en la segunda capa oculta, obteniendo resultados de clasificación del 86.62% asociando las variables psicosociales de estrés en ambientes laborales (trabajar en la parte superior de una escalera, trabajar en espacios confinados y trabajar continuamente sin tomar un descanso) para reconocer el estrés en trabajadores de construcciones. Los resultados en el nivel de exactitud del modelo y el bajo porcentaje de error, demuestran la facilidad de adaptación de la estructura matemática a las variables que requieren identificación en este tipo de procesos, con el fin de realizar intervenciones preventivas a modo de predicción muy similar a los resultados obtenidos en ésta investigación.

El modelo utilizado por Mosquera et al., (2019) de k vecinos mas cercanos y máquinas de soporte vectorial (K-NN-SVM) adaptado de Zhou et al., (2013), obtuvo resultados de clasificación del 86.66% utilizando las variables psicosociales de demandas del trabajo, el cuál logró ventajas significativas con respecto a los modelos individuales, al ver la aproximación que se hace desde el manejo únicamente de las variables de demandas del trabajo, se puede observar que el modelo funcionó bastante bien, de donde se puede inferir que el uso de las demás variables psicosociales, físicas y fisiológicas y del desarrollo de un modelo como la red neuronal de tensión superficial se adaptó mejor a las características de las variables obteniendo resultados prometedores para la investigación.

Los resultados hallados por Suárez Sánchez et al., (2011) con respecto a las variables psicosociales; demandas físicas/diseño del puesto de trabajo: esfuerzo físico, asiento incomodo, superficie de trabajo inestable o desigual, molestias en alguna parte del cuerpo debido a la tensión física; posibilidades psicosociales: considera el trabajo emocionalmente exigente; horas de trabajo y tiempo para actividad familiar y demandas emocionales del trabajo sumado a otras variables de estudio, aplicando máquinas de soporte vectorial para predecir la ocurrencia de un accidente laboral, arrojó resultados de clasificación de 97,45% teniendo en cuenta que la etiqueta de clase era accidente o no accidente. Con respecto a

estos resultados, el modelo propuesto en esta investigación va dirigido a hallar una clasificación multiclase teniendo en cuenta el grado de riesgo psicosocial, nuestro modelo obtuvo resultados de 97.37% como se puede observar en la **Tabla 4 - 6**, lo que se considera un avance para la adecuada predicción del riesgo psicosocial.

El estudio llevado a cabo por (Tzeng et al., 2004b) en un grupo de 648 enfermeras teniendo en cuenta variables psicosociales asociadas (ambiente de trabajo indirecto, ambiente de trabajo directo, salario y promoción, crecimiento personal, desafío en el trabajo, interacción y retroalimentación de pacientes y familiares, estilo de liderazgo, ambiente de trabajo y apoyo familiar y religión, satisfacción laboral general, satisfacción general con roles profesionales y felicidad laboral general) en el estudio con el objetivo de predecir la intención de abandonar su puesto de trabajo y cambiar de lugar de trabajo obtuvo resultados del orden de 89.2% utilizando máquinas de soporte vectorial, para un problema de clase binario, en este orden de ideas el algoritmo desarrollado bajo el modelo de red neuronal de tensión superficial muestra grandes ventajas para la clasificación multiclase.

Si bien las redes neuronales y las máquinas de vectores de soporte son modelos lineales que ofrecen enfoques adecuados, la aplicación potencial de un concepto físico al modelo neuronal se adapta mejor al tipo de datos presentes en las evaluaciones psicosociales. Sus posibles soluciones más rápidas permiten una mejor interpretación del modelo, así como una clasificación adecuada y tasas de error reducidas para la toma de decisiones. Como tal, el modelo de red neuronal de tensión superficial, a través del procedimiento de transformación y adecuación neuronal, puede descubrir soluciones mejoradas. En estos experimentos, como se puede observar en la **Tabla 4 - 6**, el modelo de tensión superficial superó los modelos lineales de máquinas de vectores, redes neuronales, modelos probabilísticos y modelos de árbol de decisión. La dimensionalidad y la transformación matemática redujeron 20 variables a dos variables y aumentaron el número de valores predictivos positivos y las tasas de error fueron más bajas.

La **Tabla 4 - 6**, muestra que el modelo basado en la tensión superficial física es competitivo frente a otros modelos. Proporciona resultados precisos, que son bastante similares a los

encontrados en el mundo real. Los resultados de este experimento respaldan la opinión de que la estrategia de combinar diferentes métodos (teoría de la tensión superficial física y redes neuronales artificiales) puede mejorar el rendimiento predictivo que se perseguía para el nivel de riesgo psicosocial obteniendo una exactitud en la clasificación de 97.31% en promedio y del 98.85% para el nivel de riesgo 1.

Estrategias similares se han aplicado previamente a problemas de clasificación de Seguridad y Salud Laboral, a través de la utilización de técnicas de electrocardiograma y respuesta galvánica de la piel para identificar los niveles de estrés en individuos utilizando máquinas de soporte vectorial y k-vecinos más cercanos, obteniendo resultados de clasificación alrededor de 92.75% (Sriramprakash et al., 2017). Resultados similares se han obtenido a través de la utilización de electroencefalograma para poder identificar niveles de estrés mental asociado a situaciones de control de la situación por parte de los sujetos de estudio, inducido a través de algunos mensajes que aparecían en el estudio para ver la reacción de los participantes y medir sus niveles de estrés, los algoritmos utilizados para esta clasificación fueron Naive Bayes y máquinas de soporte vectorial logrando resultados de 94% y 94.6% respectivamente (Subhani et al., 2017).

Estos resultados muestran que la aproximación realizada a través de las medidas fisiológicas en pacientes que declaraban haber tenido algún tipo de afección cardíaca y osteomuscular y a los cuales se les realizó la prueba para verificar si realmente tenían alguna relación con el nivel de riesgo psicosocial, fue una buena medida al ser incluidas como variables en el estudio. Estas variables al ser analizadas a través de pca y otras técnicas de reducción de la dimensionalidad, resultaron ser significativas en el reconocimiento de patrones, como ya lo había explicado en su estudio Sauter & Murphy, (1984), por tal razón se incluyeron en el estudio.

Las **Figura 4 - 1** y **Figura 4 - 2** muestran la forma en que el modelo agrupa e identifica el nivel de riesgo de cada individuo, lo que permite una clasificación rápida, en contraste con el caso en el que las 20 variables estadísticamente significativas se manejan conjuntamente. La

Figura 4 - 3 muestra que los procesos iterativos del modelo neuronal alcanzan rápidamente el número de iteraciones necesarias para establecer el modelo y poder proporcionar resultados objetivos y precisos.

La **Figura 4 - 4** muestra el área bajo la curva ROC, donde se muestran las ventajas de clasificación del algoritmo red neuronal de tensión superficial acorde a los resultados obtenidos en la matriz de confusión, detallando cada uno de los valores para la clasificación de cada riesgo.

Las teorías fundamentales que han servido para el desarrollo del estudio del riesgo psicosocial a nivel mundial y en Colombia (el modelo de demanda, control, apoyo social; el modelo de desequilibrio esfuerzo-recompensa y el modelo dinámico de los factores de riesgo psicosocial (Karasek, 1979; Siegrist, 1996; Villalobos, 2005) a pesar de utilizar gran cantidad de variables que abarquen todo el espectro de la organización empresarial y las variables que desde la Seguridad y la Salud en el Trabajo están inmersas en los ambientes laborales y que posiblemente sirven para identificar si existen factores de riesgo psicosocial que afectan de forma negativa al trabajador son importantes estadísticamente o representan de forma estadística una variable con gran valor en el estudio, pues se observa que muchas de ellas (en el caso docente) son solo variables generadoras de ruido (outliers).

También, se debe tener en cuenta que al momento de tomar datos para construir el conjunto de datos de riesgo psicosocial y el instrumento tenga una validez interna, la representatividad de la muestra y los sujetos de estudio pertenezcan a un mismo sector económico en éste caso el sector educativo público de colegios, para que el nivel de confianza del conjunto de datos (Dataset) pueda tener la menor cantidad de desviación en los datos. Por lo tanto, es fundamental al momento de realizar la recolección de datos futura, hacerlo por subsectores económicos, para no afectar la confiabilidad.

Finalmente se pudo comprobar la hipótesis de investigación. Se diseñó y se desarrolló un sistema de clasificación basado en técnicas inteligentes soportado en las variables psicosociales, capaz de reconocer patrones y clasificar el grado de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia, con una efectividad igual o superior al 95%.

VI. 5. Conclusiones, futuras líneas de investigación y limitaciones

5.1. Conclusiones

Conclusiones de la hipótesis

- *Un sistema de clasificación basado en técnicas inteligentes soportado en las variables psicosociales es capaz de medir el grado de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia, con una efectividad igual o superior comparado con las técnicas de medición psicosocial vigentes en Colombia.*

El diseño empleado en la metodología propuesta se logró mediante la adaptación de una teoría física, la tensión superficial a una red neuronal artificial, que incluyó variables y criterios para modelar los factores psicosociales presentes en los entornos laborales. Utilizando las bases teóricas y conceptuales se describió la metodología con un algoritmo de red neuronal de retropropagación de tensión superficial; finalmente, se validó la metodología con una prueba piloto aplicada en el sector educativo.

La experimentación se realizó con una interfaz que posibilitó el uso por parte del psicólogo organizacional. El colegio de validación fue seleccionado gracias a la participación voluntaria de los docentes; las características establecidas fueron similares debido a que se encuentra ubicado en la misma región. Se obtuvieron resultados satisfactorios que

contribuyeron a comparar el desempeño de la metodología y su aplicación en el sector educativo público. Lo que permitió comprobar la hipótesis general de investigación.

Conclusiones del objetivo general:

- *Diseñar y desarrollar un sistema de clasificación basado en técnicas inteligentes para medir el grado de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia con una efectividad igual o superior a las técnicas actualmente existentes.*

Se analizó, investigó acorde a las historias clínicas de docentes se propone el uso de 5 variables osteomusculares que ocasionan mayor ausentismo laboral y repercuten de forma psicosocial en la vida laboral del docente, para un total de 20 variables estudiadas.

En este proceso en algunos casos se hizo necesario el valorar al docente con pruebas de electromiografía, es de anotar, que una de las mayores dificultades encontradas en el trabajo al momento de recolectar este tipo de datos fue la participación voluntaria del profesorado, debido a cuestiones de no aceptar el que se adaptaran los sensores en el cuerpo.

El uso de técnicas de correlación y análisis de PCA permitió comprobar que variables tenían un alto grado de dependencia estadística, hasta obtener un conjunto de variables que estadísticamente fueran significativas para el modelo.

Las variables seleccionadas son programadas en el modelo de red neuronal de tensión superficial que permite la identificación del grado de riesgo psicosocial. Finalmente, el método, define una nueva metodología para la predicción del grado de riesgo psicosocial con una efectividad del 97.31%.

Conclusiones de cada objetivo específico:

- *Analizar los sistemas de clasificación basados en técnicas inteligentes y no inteligentes que sirvan para diseñar modelos para la identificación, evaluación y control de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia.*

Mediante una revisión bibliográfica, se seleccionaron y analizaron las metodologías, métodos, técnicas, algoritmos y variables para diseñar y desarrollar el algoritmo de clasificación de riesgo psicosocial. Lo anterior permitió identificar los modelos que han realizado aproximaciones y el uso de variables que los autores no han manejado en otras investigaciones; lo que permitió identificar el vacío en el conocimiento. Igualmente, a través de éste proceso se pudo establecer las bases teóricas y conceptuales del objeto de estudio, desarrollo del método propuesto y su experimentación.

Se analizaron los referentes teóricos del estudio psicosocial utilizados para la construcción de la batería de riesgos psicosocial, instrumento avalado en Colombia para el estudio de los factores psicosociales y se identificó las variables relevantes en éste tipo de estudios según lo planteado en el instrumento.

- *Definir las variables independientes (nuevas y existentes) en la identificación de los riesgos psicosociales y seleccionar las variables más efectivas mediante técnicas informáticas y/o estadísticas, con el fin de lograr un mayor nivel de automatización en el proceso de identificación de este riesgo.*

Se llevó acabo la evaluación de riesgo psicosocial aplicando la batería de riesgo psicosocial Intralaboral en docentes de colegios públicos de municipios pertenecientes al área metropolitana de la ciudad de Cali, revisando los antecedentes médicos y de ausentismo laboral para evaluar las nuevas variables incluidas en el estudio (Según refiere (Sauter & Murphy, 1984) son síntomas reconocidos como afecciones somáticas crónicas

asociadas a la existencia de exposición a factores de riesgo psicosocial. Se realizó varios procedimientos estadísticos como: Preprocesamiento de datos, pruebas de correlación estadística de pearson, reducción de la dimensionalidad a través de algoritmos genéticos, PCA, gain ratio attribute eval (Mosquera, Castrillón, et al., 2018b) y algoritmo ReliefF. Estos procedimientos de significancia estadística, permitieron reducir la cantidad de variables hasta un número importante que permitiera una adecuada clasificación gracias a la eliminación de datos que se consideran outliers (datos atípicos dentro del conjunto de datos). En este trabajo de investigación se logró por medio de esta metodología descartar 103 características psicosociales y encontrar 12 características que probaron ser efectivas en la identificación del grado de riesgo psicosocial.

- *Desarrollar un sistema de clasificación basado en técnicas inteligentes que permita identificar y medir los factores de riesgo psicosocial en docentes de educación básica primaria y secundaria en colegios públicos de Colombia.*

Principalmente se investigaron dos métodos de clasificación: máquinas de soporte vectorial y redes neuronales, y se investigó un método con el cual emular el grado de riesgo psicosocial: la tensión superficial física.

Como resultado de lo anterior, se utilizaron las redes neuronales de retropropagación y se modifica una de las capas introduciendo la teoría de tensión superficial, con la cual se logra reducir la dimensionalidad de las variables en la red adaptando los dos conceptos de perímetro y masa a las variables psicosociales que se introducen modificando los pesos y reemplazando estos valores por las medidas de tensión superficial de cuatro líquidos, los cuales emulan los cuatro grados de riesgo psicosocial. Finalmente, en este punto se obtienen los resultados propuestos de innovación y mejoramiento del proceso de clasificación.

- *Validar el sistema de clasificación propuesto con el fin de establecer su efectividad en comparación con los métodos tradicionalmente empleados por los especialistas*

aplicados en docentes de educación básica primaria y secundaria en colegios públicos de Colombia.

La metodología propuesta está diseñada de acuerdo a las características de los colegios públicos en Colombia, lo cual permite que pueda ser aplicada en diferentes instituciones educativas cuya jornada sea continua (mañana, tarde), empleando la metodología propuesta con un conjunto de datos acorde a cada región, con el fin de obtener un mejor análisis sobre su verdadero impacto en la disminución en las tasas morbilidad y de ausentismo de los docentes.

5.2. Futuras líneas de investigación

A partir de este estudio, se puede vislumbrar futuros estudios a realizar en el campo de la Inteligencia Artificial, es interesante, la posibilidad que ofrecen los sistemas inteligentes de construir y evaluar un clasificador aplicando técnicas inteligentes a la identificación de riesgo psicosocial como herramienta para la gestión y control gerencial de los sistemas de Seguridad y Salud en el Trabajo, que sirva como elemento para fortalecer la toma de decisiones en la política de la alta dirección empresarial.

La aplicación de la metodología predictiva de red neuronal de tensión superficial, a los factores de riesgo psicosocial de instituciones educativa públicas, muestra que, se abre una línea de trabajo y de investigación que permite apoyar la labor de los grupos encargados de la Seguridad y Salud en el Trabajo y facilita la labor de la identificación, prevención, promoción de la salud, en los individuos claves que están expuestos a factores de riesgo y cuyo resultado de clasificación arroja valores de Alto y Muy Alto y, permite la identificación, reconocimiento y clasificación de los riesgos psicosociales en los entornos de trabajo.

Este estudio muestra las virtudes del desarrollo del modelo con un algoritmo aplicado de teoría física. Este enfoque es una herramienta poderosa para la prevención y el tratamiento del riesgo psicosocial en docentes de colegios públicos. Es necesario recopilar conjuntos

de datos por diferentes áreas geográficas, incluidas las variables ergonómicas. También es importante definir que existe una diferencia entre el estrés laboral, la satisfacción laboral y el bienestar psicosocial, para que la investigación futura desarrolle aplicaciones en el estrés laboral y la satisfacción laboral con algoritmos de inteligencia artificial.

También, se deja claro la necesidad de realizar mejoramiento continuo a las herramientas que tiene el país para la identificación, evaluación, control y seguimiento que se hace a los factores de riesgo psicosocial, tanto en empresas públicas como privadas.

Como vemos las disciplinas relacionadas con las técnicas de machine learning, incluyen el estudio psicológico del aprendizaje humano, el estudio de la evolución, la teoría del control adaptativo, el estudio de las prácticas educativas, la neurociencia, el comportamiento organizacional y la economía. Aunque la última década ha visto un mayor crecimiento con estos otros campos, recién estamos comenzando a aprovechar las posibles sinergias y la diversidad de formalismos y métodos experimentales utilizados en estos campos múltiples para estudiar sistemas que mejoran con la experiencia (Jordan & Mitchell, 2015), como son los psicosociales y organizacionales que abren un espacio muy importante en la prevención de riesgos laborales y en procesos asociados a clima laboral y cultura de entornos saludables.

Se necesitan también la intervención de otras técnicas de la inteligencia artificial, como la visión artificial, con la cual se logre desarrollar la identificación de ciertos comportamientos, gestos, o estados de ánimo, que permitan reconocer patrones y clasificar problemas de índole psicosocial en los ambientes laborales, permitiendo anticiparse a través de la prevención a los posibles estados de afección que puedan desarrollar los trabajadores, no solo del subsector educativo, sino, de los demás sectores de la economía.

5.3. Limitaciones

Para éste proyecto el poder conseguir la información de los pacientes de forma coordinada, el poder cablear adecuadamente a los docentes y el tiempo de labor docente, fue una de

las mayores restricciones que se tuvo al momento de adelantar la consecución de los datos y en el que mas tiempo se invirtió (Un año y ocho meses).

VII. Referencias

- Abma, F. I., Brouwer, S., de Vries, H. J., Arends, I., Robroek, S. J., Cuijpers, M. P., van der Wilt, G. J., Bültmann, U., & van der Klink, J. J. (2016). The capability set for work: Development and validation of a new questionnaire. *Scandinavian Journal of Work, Environment & Health*, 42(1), 34–42. <https://doi.org/10.5271/sjweh.3532>
- Adamson, A. W., & Gast, A. P. (1967). Arthur W. Adamson, Alice P. Gast—Physical chemistry of surfaces. *New York: Interscience Publishers*, 150(180).
- Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37, 13–49. <https://doi.org/10.1016/j.tele.2019.01.007>
- Aliabadi, M. (2015). Prediction of hearing loss among the noise-exposed workers in a steel factory using artificial intelligence approach. *Int Arch Occup Environ Health*, 88(6), 779–787.
- Alpaydin, E. (2020). *Introduction to Machine Learning* (Fourth). The MIT Press.
- Al-Shargie, F. (2019). *Multilevel Assessment of Mental Stress using SVM with ECOC: An EEG Approach* [Preprint]. engrXiv. <https://doi.org/10.31224/osf.io/7v9ks>
- Anand, R. S., & Kumar, V. (2018). EEG-metric based mental stress detection. *Network Biology*, 8(1), 25–34.
- Andayani, U., Nababan, E. B., Siregar, B., Muchtar, M. A., Nasution, T. H., & Siregar, I. (2017). Optimization backpropagation algorithm based on Nguyen-Widrom adaptive weight and adaptive learning rate. *2017 4th International Conference on*

- Industrial Engineering and Applications (ICIEA)*, 363–367.
<https://doi.org/10.1109/IEA.2017.7939239>
- Ansoleaga, E., Portales, U. D., Toro, J. P., & Portales, U. D. (2014). Mental Health and Nature of Work: When the Emotional Demands Becomes Inevitable. *Revista Psicologia: Organizações e Trabalho*, 14(2), 180–189.
- Baradaran, V., Ghadami, S., & Malihi, S. E. (2008). A Multi Objective approach for selecting solutions to improve job satisfaction an empirical case analysis. 2008 *IEEE International Conference on Industrial Engineering and Engineering Management*, 1945–1948. <https://doi.org/10.1109/IEEM.2008.4738211>
- Bauer, G. F., & Hämmig, O. (2014). *Bridging Occupational, Organizational and Public Health*. Springer Netherlands. <https://doi.org/10.1007/978-94-007-5640-3>
- Belkin, M., Hsu, D., Ma, S., & Mandal, S. (2019). Reconciling modern machine learning practice and the bias-variance trade-off. *ArXiv:1812.11118 [Cs, Stat]*.
<http://arxiv.org/abs/1812.11118>
- Benítez, R., Escudero, G., & Kanaan, S. (2014). *Inteligencia artificial avanzada* (2da ed.). UOC.
- Berens, J., Schneider, K., Görtz, S., Oster, S., & Burghoff, J. (2019). Early Detection of Students at Risk—Predicting Student Dropouts Using Administrative Student Data from German Universities and Machine Learning Methods. *Journal of Educational Data Mining*, 11(3), 1–41.
- Berry, M. V. (1971). The molecular mechanism of surface tension. *Physics Education*, 6(2). <https://doi.org/10.1088/0031-9120/6/2/001>
- Bitalino. (2017). *BITalino* [OpenSignals]. <http://www.bitalino.com/>

- Botero Alvarez, C. C. (2013). Riesgo psicosocial intralaboral y “burnout” en docentes universitarios de algunos países latinoamericanos. *Cuadernos de Administración*, 28(48), 118–133. <https://doi.org/10.25100/cdea.v28i48.460>
- British Standards Institution (BSI). (2011). *PAS1010: Guidance on the management of psychosocial risks in the workplace*. BSI.
- Brown, R. C. (1947). The fundamental concepts concerning surface tension and capillarity. *Proceedings of the Physical Society*, 59(3), 429–448. <https://doi.org/10.1088/0959-5309/59/3/310>
- Bruhn, A., & Frick, K. (2011). Why it was so difficult to develop new methods to inspect work organization and psychosocial risks in Sweden. *Safety Science*, 49(4), 575–581. <https://doi.org/10.1016/j.ssci.2010.07.011>
- Burke, R. J., & Pignata, S. (2020). *Handbook of Research on Stress and Well-Being in the Public Sector* (First, Vol. 1). Edward Elgar.
- Canadian Standards Association. (2013). *Psychological health and safety in the workplace—Prevention, promotion, and guidance to staged implementation*. Bureau de normalisation du Québec.
- Cardenas Gonzalo, D. (2015). Influencia de los síntomas físicos sobre el estrés laboral y familiar. *Dyna Management*, 3(3), 248–262. <https://doi.org/10.6036/MN7831>
- Chakraborty, U. (2020). *Artificial Intelligence for All: Transforming Every Aspect of Our Life* (1a ed.). Bpb publications.
- Chen, L., Pan, X., Zhang, Y.-H., Liu, M., Huang, T., & Cai, Y.-D. (2019). Classification of Widely and Rarely Expressed Genes with Recurrent Neural Network. *Computational and Structural Biotechnology Journal*, 17, 49–60. <https://doi.org/10.1016/j.csbj.2018.12.002>
- Cheng, J., Wang, L., & Xiong, Y. (2019). Ensemble of cuckoo search variants. *Computers & Industrial Engineering*, 135, 299–313. <https://doi.org/10.1016/j.cie.2019.06.015>

- Cox, T., Griffiths, A., Rial-González, E., & Agencia Europea para la Seguridad y la Salud en el Trabajo. (2006). *Investigación sobre el estrés relacionado con el trabajo*. Oficina de publicaciones oficiales de las Comunidades europeas.
- Cuixart, C. N. (2012). *NTP:926 Factores psicosociales: Metodología de evaluación*. Centro Nacional de Condiciones del Trabajo.
<https://www.insst.es/documents/94886/326879/926w.pdf/cdecbd91-70e8-4cac-b353-9ea39340e699>
- Czaja, S. J., & Nair, S. N. (2006). Human Factors Engineering and Systems Design. En G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (pp. 32–49). John Wiley & Sons, Inc. <https://doi.org/10.1002/0470048204.ch2>
- Darvishi, E., Khotanlou, H., Khoubi, J., Giahi, O., & Mahdavi, N. (2017). Prediction Effects of Personal, Psychosocial, and Occupational Risk Factors on Low Back Pain Severity Using Artificial Neural Networks Approach in Industrial Workers. *Journal of Manipulative and Physiological Therapeutics*, 40(7), 486–493.
<https://doi.org/10.1016/j.jmpt.2017.03.012>
- Dediu, V., Leka, S., & Jain, A. (2018). Job demands, job resources and innovative work behaviour: A European Union study. *European Journal of Work and Organizational Psychology*, 27(3), 310–323.
<https://doi.org/10.1080/1359432X.2018.1444604>
- Dimsdale, J. E. (2019). *Trastorno de síntomas somáticos* [Profesional]. Manual MSD.
<https://www.msmanuals.com/es/professional/trastornos-psiqui%C3%A1tricos/trastornos-de-s%C3%ADntomas-som%C3%A1ticos-y-relacionados/trastorno-de-s%C3%ADntomas-som%C3%A1ticos>
- Dollard, M. F., & Neser, D. Y. (2013). Worker health is good for the economy: Union density and psychosocial safety climate as determinants of country differences in

- worker health and productivity in 31 European countries. *Social Science & Medicine*, 92, 114–123. <https://doi.org/10.1016/j.socscimed.2013.04.028>
- Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (Second). John Wiley & Sons.
- El Yafrani, M., & Ahiod, B. (2018). Efficiently solving the Traveling Thief Problem using hill climbing and simulated annealing. *Information Sciences*, 432, 231–244. <https://doi.org/10.1016/j.ins.2017.12.011>
- El-Batawi, M. A. (1988). Psychosocial health problems of workers in developing countries. En *Psychosocial factors at work and their relationship with health* (1a ed., Vol. 1, pp. 15–20). World Health Organization.
- Elhoone, H., Zhang, T., Anwar, M., & Desai, S. (2020). Cyber-based design for additive manufacturing using artificial neural networks for Industry 4.0. *International Journal of Production Research*, 58(9), 2841–2861. <https://doi.org/10.1080/00207543.2019.1671627>
- Esling, P., & Devis, N. (2020). Creativity in the era of artificial intelligence. *ArXiv:2008.05959 [Cs]*. <http://arxiv.org/abs/2008.05959>
- Espinoza, L. M., & Villalobos, D. G. (2015). *Prevalencia de riesgo psicosocial en un grupo de docentes y directivos del distrito capital* [Tesis de Maestría]. Universidad del Rosario.
- European agency for safety and health and work. (2011). *Annual_report_2010_summary_es.pdf*. EU-Osha 2010. https://osha.europa.eu/es/publications/annual_report/ar_summary_2010/view
- European Agency for Safety and Health at Work. (2012). Drivers and barriers for psychosocial risk management: An analysis of the findings of the European Survey of Enterprises on New and Emerging Risks (ESENER). *report. Publ. Office of the Europ. Union*.

- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Fecode. (2015). *Acta-de-acuerdos-fecode_MEN-Mayo-7-2015*.
<https://fecode.edu.co/index.php/actas-de-acuerdo>
- Fernandes de Mello, R., & Antonelli Ponti, M. (2018). *Machine Learning: A Practical Approach on the Statistical Learning Theory*. Springer International Publishing.
<https://doi.org/10.1007/978-3-319-94989-5>
- Fowkes, F. M. (1961). *Determination of interfacial tensions, contact angles, and dispersion forces in surfaces*. 66, 1.
- Frone, M. R., & Tidwell, M.-C. O. (2015). The meaning and measurement of work fatigue: Development and evaluation of the Three-Dimensional Work Fatigue Inventory (3D-WFI). *Journal of Occupational Health Psychology*, 20(3), 273–288.
<https://doi.org/10.1037/a0038700>
- Galatzer-Levy, I. R., Karstoft, K.-I., Statnikov, A., & Shalev, A. Y. (2014). Quantitative forecasting of PTSD from early trauma responses: A Machine Learning application. *Journal of Psychiatric Research*, 59, 68–76.
<https://doi.org/10.1016/j.jpsychires.2014.08.017>
- García Castro, C. M., & Muñoz Sánchez, A. I. (2013). Health and work of district schools faculty of the district one of Bogotá. *AVANCES EN ENFERMERÍA*, 13.
- García-Herrero, S., Lopez-Garcia, J. R., Herrera, S., Fontaneda, I., Báscones, S. M., & Mariscal, M. A. (2017). The Influence of Recognition and Social Support on European Health Professionals' Occupational Stress: A Demands-Control-Social Support-Recognition Bayesian Network Model. *BioMed Research International*, 2017, 1–14. <https://doi.org/10.1155/2017/4673047>

- González Fuentes, A., Busto Serrano, N. M., Sánchez Lasheras, F., Fidalgo Valverde, G., & Suárez Sánchez, A. (2020). Prediction of Health-Related Leave Days among Workers in the Energy Sector by Means of Genetic Algorithms. *Energies*, *13*(10), 2475. <https://doi.org/10.3390/en13102475>
- González Herranz, R. (2016). *Sistemas de ayuda al diagnóstico y a la terapia funcional en enfermedades neurodegenerativas* [Disertación doctoral]. Universidad Politécnica de Madrid.
- Hadi, W., El-Khalili, N., AlNashashibi, M., Issa, G., & AlBanna, A. A. (2019). Application of data mining algorithms for improving stress prediction of automobile drivers: A case study in Jordan. *Computers in Biology and Medicine*, *114*, 103474. <https://doi.org/10.1016/j.compbiomed.2019.103474>
- Hallner, D., & Hasenbring, M. (2004). Classification of psychosocial risk factors (yellow flags) for the development of chronic low back and leg pain using artificial neural network. *Neuroscience Letters*, *361*(1–3), 151–154. <https://doi.org/10.1016/j.neulet.2003.12.107>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed). Morgan Kaufmann Publishers.
- Hegde, J., & Rokseth, B. (2020). Applications of machine learning methods for engineering risk assessment – A review. *Safety Science*, *122*, 104492. <https://doi.org/10.1016/j.ssci.2019.09.015>
- Hernández, M. G. G., Martínez, R. M. R., Maldonado-Macias, A. A., & Leal, J. S. (Eds.). (2020). *Work Stress and Psychosocial Factors in the Manufacturing Industry: A Literature Review* (First, Vol. 1). IGI Global. <https://doi.org/10.4018/978-1-7998-1052-0>

- Hesser, D. F., & Markert, B. (2019). Tool wear monitoring of a retrofitted CNC milling machine using artificial neural networks. *Manufacturing Letters*, *19*, 1–4.
<https://doi.org/10.1016/j.mfglet.2018.11.001>
- Holte, R. C. (1993). Very Simple Classification Rules Perform Well on Most Commonly Used Datasets. *Machine Learning*, *11*, 63–91.
- Houtman, I., van Zwieten, M., Leka, S., Jain, A., & de Vroome, E. (2020). Social Dialogue and Psychosocial Risk Management: Added Value of Manager and Employee Representative Agreement in Risk Perception and Awareness. *International Journal of Environmental Research and Public Health*, *17*(10), 3672.
<https://doi.org/10.3390/ijerph17103672>
- Iavicoli, S., Leka, S., Jain, A., Persechino, B., Rondinone, B. M., Ronchetti, M., & Valenti, A. (2014). Hard and soft law approaches to addressing psychosocial risks in Europe: Lessons learned in the development of the Italian approach. *Journal of Risk Research*, *17*(7), 855–869. <https://doi.org/10.1080/13669877.2013.822911>
- Jacobson, S. H., & Yücesan, E. (2004). Analyzing the Performance of Generalized Hill Climbing Algorithms. *Journal of Heuristics*, *10*(4), 387–405.
<https://doi.org/10.1023/B:HEUR.0000034712.48917.a9>
- Jain, A., Dediu, V., Zwetsloot, G., & Leka, S. (2017). Workplace Innovation and Wellbeing at Work: A Review of Evidence and Future Research Agenda. En P. Oeij, D. Rus, & F. D. Pot (Eds.), *Workplace Innovation* (pp. 111–128). Springer International Publishing. https://doi.org/10.1007/978-3-319-56333-6_8
- Jasper, J. J. (1972). *The Surface Tension of Pure Liquid Compounds*. *1*(4), 841–1010.
<https://doi.org/10.1063/1.3253106>
- Jebelli, H., Khalili, M. M., & Lee, S. (2019). Mobile EEG-Based Workers' Stress Recognition by Applying Deep Neural Network. En I. Mutis & T. Hartmann (Eds.),

- Advances in Informatics and Computing in Civil and Construction Engineering* (pp. 173–180). Springer International Publishing. https://doi.org/10.1007/978-3-030-00220-6_21
- Johnson, J. V., & Hall, E. M. (1988). Job strain, work place social support, and cardiovascular disease: A cross-sectional study of a random sample of the Swedish working population. *American Journal of Public Health*, 78(10), 1336–1342.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- Jung, J. S., Park, S. J., Kim, E. Y., Na, K.-S., Kim, Y. J., & Kim, K. G. (2019). Prediction models for high risk of suicide in Korean adolescents using machine learning techniques. *PLOS ONE*, 14(6), e0217639. <https://doi.org/10.1371/journal.pone.0217639>
- Karasek, R. A. (1979). Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. *Administrative Science Quarterly*, 24(2), 285. <https://doi.org/10.2307/2392498>
- Kataoka, H., Kano, H., Yoshida, H., Saijo, A., Yasuda, M., & Osumi, M. (1998). Development of a skin temperature measuring system for non-contact stress evaluation. *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286)*, 2, 940–943. <https://doi.org/10.1109/IEMBS.1998.745598>
- Kato, E. R. R., Aranha, G. D. de A., & Tsunaki, R. H. (2018). A new approach to solve the flexible job shop problem based on a hybrid particle swarm optimization and Random-Restart Hill Climbing. *Computers & Industrial Engineering*, 125, 178–189. <https://doi.org/10.1016/j.cie.2018.08.022>

- Kohonen, T. (1989). *Self-Organization and Associative Memory* (Vol. 8). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-88163-3>
- Kortum, E., & Leka, S. (2014). Tackling psychosocial risks and work-related stress in developing countries: The need for a multilevel intervention framework. *International Journal of Stress Management*, 21(1), 7–26. <https://doi.org/10.1037/a0035033>
- Kortum, E., Leka, S., & Cox, T. (2010). Psychosocial risks and work-related stress in developing countries: Health impact, priorities, barriers and solutions. *International Journal of Occupational Medicine and Environmental Health*, 23(3). <https://doi.org/10.2478/v10001-010-0024-5>
- Kubat, M. (2017). *An Introduction to Machine Learning*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-63913-0>
- Lahmiri, S., & Bekiros, S. (2019). Can machine learning approaches predict corporate bankruptcy? Evidence from a qualitative experimental design. *Quantitative Finance*, 19(9), 1569–1577. <https://doi.org/10.1080/14697688.2019.1588468>
- Lal, A., Dhinesh, R., Harish, C. S., Moorthy, D. M. J., & Madhumathi, R. (2018). Stress Detection using Machine Learning. *International Journal for Trends in Engineering & Technology*, 28(1), 37–39.
- Langenhan, M. K., Leka, S., & Jain, A. (2013). Psychosocial Risks: Is Risk Management Strategic Enough in Business and Policy Making? *Safety and Health at Work*, 4(2), 87–94. <https://doi.org/10.1016/j.shaw.2013.04.003>
- Larrabee, J. H., Janney, M. A., Ostrow, C. L., Withrow, M. L., Hobbs, G. R., & Burant, C. (2003). Predicting Registered Nurse Job Satisfaction and Intent to Leave: *JONA: The Journal of Nursing Administration*, 33(5), 271–283. <https://doi.org/10.1097/00005110-200305000-00003>

- Law, K.-Y., & Zhao, H. (2016). *Surface wetting: Characterization, contact angle, and fundamentals*. (2a ed.). Springer International Publishing.
- Lawrence, S. A., Jordan, P. J., & Callan, V. J. (2015). Initial validation of the support mobilization for work stressors inventory. *Australian Journal of Management*, 40(4), 587–612. <https://doi.org/10.1177/0312896214528186>
- Leka, S., & Cox, T. (2011). *Guide Prima-EF. Guidance on the European Framework for Psychosocial Risk Management. A Resource for Employers and Workers Representatives*. WHO.
- Leka, Stavroula, Jain, A., & Lerouge, L. (2017). Work-Related Psychosocial Risks: Key Definitions and an Overview of the Policy Context in Europe. En L. Lerouge (Ed.), *Psychosocial Risks in Labour and Social Security Law* (pp. 1–12). Springer International Publishing. https://doi.org/10.1007/978-3-319-63065-6_1
- Leka, Stavroula, Van Wassenhove, W., & Jain, A. (2015). Is psychosocial risk prevention possible? Deconstructing common presumptions. *Safety Science*, 71, 61–67. <https://doi.org/10.1016/j.ssci.2014.03.014>
- Li, B., Mendenhall, J., & Meiler, J. (2019). Interfaces Between Alpha-helical Integral Membrane Proteins: Characterization, Prediction, and Docking. *Computational and Structural Biotechnology Journal*, 17, 699–711. <https://doi.org/10.1016/j.csbj.2019.05.005>
- Li, L., Cen, Z.-Y., Tseng, M.-L., Shen, Q., & Ali, M. H. (2021). Improving short-term wind power prediction using hybrid improved cuckoo search arithmetic—Support vector regression machine. *Journal of Cleaner Production*, 279, 123739. <https://doi.org/10.1016/j.jclepro.2020.123739>
- Li, Z., Sun, D., Zhu, R., & Lin, Z. (2017). Detecting event-related changes in organizational networks using optimized neural network models. *PLOS ONE*, 12(11), e0188733. <https://doi.org/10.1371/journal.pone.0188733>

- Lippel, K., & Quinlan, M. (2011). Regulation of psychosocial risk factors at work: An international overview. *Safety Science*, 49(4), 543–546.
<https://doi.org/10.1016/j.ssci.2010.09.015>
- Lotfan, S., Shahyad, S., Khosrowabadi, R., Mohammadi, A., & Hatef, B. (2019). Support vector machine classification of brain states exposed to social stress test using EEG-based brain network measures. *Biocybernetics and Biomedical Engineering*, 39(1), 199–213. <https://doi.org/10.1016/j.bbe.2018.10.008>
- Loy-Benitez, J., Heo, S., & Yoo, C. (2020). Soft sensor validation for monitoring and resilient control of sequential subway indoor air quality through memory-gated recurrent neural networks-based autoencoders. *Control Engineering Practice*, 97, 104330. <https://doi.org/10.1016/j.conengprac.2020.104330>
- Macleod, D. B. (1923). On a relation between surface tension and density. *Transactions of the Faraday Society*, 19(July), 38. <https://doi.org/10.1039/tf9231900038>
- Maqsoom, A., Mughees, A., Zahoor, H., Nawaz, A., & Mazher, K. M. (2020). Extrinsic psychosocial stressors and workers' productivity: Impact of employee age and industry experience. *Applied Economics*, 52(26), 2807–2820.
<https://doi.org/10.1080/00036846.2019.1696936>
- Márquez Gómez, M. (2020). Prediction of work-related musculoskeletal discomfort in the meat processing industry using statistical models. *International Journal of Industrial Ergonomics*, 75, 102876. <https://doi.org/10.1016/j.ergon.2019.102876>
- Maynard, M. (2020). *Maynard, Morgan—Neural Networks_ Introduction to Artificial Neurons, Backpropagation and Multilayer Feedforward Neural Networks with Real-World Applications (Advance. Maynard.*
- McClelland, J. L., & Rumelhart, E. (eds.). (1986). *Parallel Distributed Processing (Vol. 2)*. MIT Press.

- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115–133.
<https://doi.org/10.1007/BF02478259>
- Mehanović, D., Mašetić, Z., & Kečo, D. (2020). Prediction of Heart Diseases Using Majority Voting Ensemble Method. En A. Badnjevic, R. Škrbić, & L. Gurbeta Pokvić (Eds.), *CMBEBIH 2019* (Vol. 73, pp. 491–498). Springer International Publishing. https://doi.org/10.1007/978-3-030-17971-7_73
- Memish, K., Martin, A., Bartlett, L., Dawkins, S., & Sanderson, K. (2017). Workplace mental health: An international review of guidelines. *Preventive Medicine*, 101, 213–222. <https://doi.org/10.1016/j.ypmed.2017.03.017>
- Meneses, A. A. de M., da Silva, P. V., Nast, F. N., Araujo, L. M., & Schirru, R. (2020). Application of Cuckoo Search algorithm to Loading Pattern Optimization problems. *Annals of Nuclear Energy*, 139, 107214.
<https://doi.org/10.1016/j.anucene.2019.107214>
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. *Advances in Neural Information Processing Systems*, 3111–3119.
- Ministerio de Educación. (2015). *Decreto_1655_de_2015.pdf*. Min_educación.
- Ministerio de la Protección Social. (2008). *Resolución número 002646 de 2008*.
<http://www.saludcapital.gov.co/Documentos%20Salud%20Ocupacional/RESOL.%202646%20DE%202008%20RIESGO%20PSICOSOCIAL.pdf>
- Ministerio del Trabajo. (2014). *Decreto_1477_del_5_de_agosto_de_2014*.
https://www.mintrabajo.gov.co/documents/20147/36482/decreto_1477_del_5_de_agosto_de_2014.pdf/b526be63-28ee-8a0d-9014-8b5d7b299500
- Ministerio del trabajo. (2015). *Guía técnica general para la Promoción, prevención e intervención de los factorespsicosociales y sus efectos en población trabajadora*.

- Javegraf. <http://fondoriesgoslaborales.gov.co/wp-content/uploads/2018/09/01-Guia-tecnica-general.pdf>
- Ministerio del Trabajo, & Organización Iberoamericana de Seguridad Social, Oiss. (2013). *II_ENCUESTA_NACIONAL_CONDICIONES_SST_COLOMBIA_2013.pdf* (Oiss).
- Moncada, S., Utzet, M., Molinero, E., Llorens, C., Moreno, N., Galtés, A., & Navarro, A. (2014). The copenhagen psychosocial questionnaire II (COPSOQ II) in Spain-A tool for psychosocial risk assessment at the workplace: Copenhagen Psychosocial Questionnaire II in Spain. *American Journal of Industrial Medicine*, 57(1), 97–107. <https://doi.org/10.1002/ajim.22238>
- Morrow, A. S., Campos Vega, A. D., Zhao, X., & Liriano, M. M. (2020). Leveraging Machine Learning to Identify Predictors of Receiving Psychosocial Treatment for Attention Deficit/Hyperactivity Disorder. *Administration and Policy in Mental Health and Mental Health Services Research*, 47(4). <https://doi.org/10.1007/s10488-020-01045-y>
- Mosquera Navarro, R. (2019). Psychosocial risk level on teachers schools V1. *IEEE Dataport*, 1(1). <http://dx.doi.org/10.21227/jdzw-7e82>
- Mosquera Navarro, R., Gómez, O. D. C., Osorio, L. P., & García, A. C. (2018). Classification system for the predicting of psychosocial risk level in public-school teachers based on Artificial Intelligence. *XVIII Conferencia de La Asociación Española Para La Inteligencia Artificial (CAEPIA)*, 1367–1372.
- Mosquera, R, Castrillón Gómez, O. D., & Parra-Osorio, L. (2019). *Algorithm based on Physical Surface Tension for the Prediction of Psychosocial-risk Level in Public School Teachers. [Source Code]. Doi.org/10.24433/CO.4268666.v1.1. (1.1) [Matlab]. Code Ocean.*

- Mosquera, Rodolfo. (2018). Psychosocial Risk Level Colombian Teachers School Repository 2016-2017. *Zenodo*, 2. <https://doi.org/10.5281/zenodo.1298610>.
- Mosquera, Rodolfo, Castrillón, O. D., & Parra, L. (2018a). Support Vector Machines, Naïve Bayes Classifier and Genetic Algorithms for the Prediction of Psychosocial Risks in Teachers of Colombian Public Schools. *Información tecnológica*, 29(6), 153–162. <https://doi.org/10.4067/S0718-07642018000600153>
- Mosquera, Rodolfo, Castrillón, O. D., & Parra, L. (2018b). Predicción de Riesgos Psicosociales en Docentes de Colegios Públicos Colombianos utilizando Técnicas de Inteligencia Artificial. *Información tecnológica*, 29(4), 267–280. <https://doi.org/10.4067/S0718-07642018000400267>
- Mosquera, Rodolfo, Castrillón, O. D., & Parra, L. (2018c). Máquinas de Soporte Vectorial, Clasificador Naïve Bayes y Algoritmos Genéticos para la Predicción de Riesgos Psicosociales en Docentes de Colegios Públicos Colombianos. *Información tecnológica*, 29(6), 153–162. <https://doi.org/10.4067/S0718-07642018000600153>
- Mosquera, Rodolfo, Castrillón, O. D., & Parra-Osorio, L. (2019). Aplicación del modelo híbrido k-nearest neighbors- Support Vector Machine para la predicción del riesgo psicosocial en docentes de colegios públicos colombianos. *Proceedings of the 17th Latin American and Caribbean Conference for Engineering and Technology*, 1, 5.
- Mosquera, Rodolfo, Gómez, O. D. C., Osorio, L. P., & García, A. C. (2018). Classification system for the predicting of psychosocial risk level in public-school teachers based on Artificial Intelligence. *Proceedings of XVIII Conferencia de La Asociación Española Para La Inteligencia Artificial*, 1, 1367–1372.
- Mosquera, Rodolfo, Parra-Osorio, L., & Castrillón, O. D. (2016). Metodología para la Predicción del Grado de Riesgo Psicosocial en Docentes de Colegios

- Colombianos utilizando Técnicas de Minería de Datos. *Información tecnológica*, 27(6), 259–272. <https://doi.org/10.4067/S0718-07642016000600026>
- Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Interpretable machine learning: Definitions, methods, and applications. *Proceedings of the National Academy of Sciences*, 116(44), 22071–22080. <https://doi.org/10.1073/pnas.1900654116>
- Navascués, G. (1979). Liquid surfaces: Theory of surface tension. *Reports on Progress in Physics*, 42(7), 59. <https://doi.org/10.1088/0034-4885/42/7/002>
- Neal, B., Mittal, S., Baratin, A., Tantia, V., Scicluna, M., Lacoste-Julien, S., & Mitliagkas, I. (2019). A Modern Take on the Bias-Variance Tradeoff in Neural Networks. *ArXiv:1810.08591 [Cs, Stat]*. <http://arxiv.org/abs/1810.08591>
- Niedhammer, I., & Chastang, J.-F. (2013). Psychosocial work factors and sickness absence in 31 countries in Europe. *European Journal of Public Health*, 23(4), 622–628.
- Olaya Arévalo, C. (2015). *Síndrome de burnout o síndrome de agotamiento profesional (sap) en el trabajo de los docentes distritales de la localidad de usme* [Tesis de Maestría].
- Pavelka, A., & Prochazka, A. (2004). Algorithms for initialization of neural network weights. *Proceedings of the 12th Annual Conference*, 453–459.
- Posada Quintero, J. I., Molano Vergara, P. N., Parra Hernández, R. M., Brito Osorio, F. Y., & Rubio Orozco, E. A. (2019). Prevalencia del Síndrome de Burnout en docentes: Factores asociados al estatuto de vinculación laboral en Colombia. *Revista Interamericana de Psicología Ocupacional*, 37(2), 119–133. <https://doi.org/10.21772/ripo.v37n2a04>

- Posada-Quintero, H. F., Molano-Vergara, P. N., Parra-Hernández, R. M., & Posada-Quintero, J. I. (2020). Analysis of Risk Factors and Symptoms of Burnout Syndrome in Colombian School Teachers under Statutes 2277 and 1278 Using Machine Learning Interpretation. *Social Sciences*, 9(3), 30.
<https://doi.org/10.3390/socsci9030030>
- Priya, A., Garg, S., & Tigga, N. P. (2020). Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms. *Procedia Computer Science*, 167, 1258–1267. <https://doi.org/10.1016/j.procs.2020.03.442>
- Rashidi, H. H., Tran, N. K., Betts, E. V., Howell, L. P., & Green, R. (2019). Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. *Academic Pathology*, 6, 237428951987308.
<https://doi.org/10.1177/2374289519873088>
- Restrepo, G. C. (2011). *Factores de riesgo psicosocial que afectan el ejercicio pedagógico de los docentes del núcleo educativo nº 8 de la ciudadela cuba de pereira* [Tesis de Maestría]. Católica de Pereira.
- Robnik-S̃, M., & Kononenko, I. (2003). Theoretical and Empirical Analysis of ReliefF and RReliefF. *Machine Learning*, 53(1–2), 23–69.
<https://doi.org/10.1023/A:1025667309714>
- Robnik-S̃, M., & Kononenko, I. (2013). Theoretical and Empirical Analysis of ReliefF and RReliefF. *Machine Learning*, 53, 23–69.
- Rodriguez, L. (2007). *Probabilidad y estadística básica para ingenieros*. Escuela Superior Politécnica del Litoral Instituto de Ciencias Matemáticas.
- Rodríguez-Arce, J., Lara-Flores, L., Portillo-Rodríguez, O., & Martínez-Méndez, R. (2020). Towards an anxiety and stress recognition system for academic environments based on physiological features. *Computer Methods and Programs in Biomedicine*, 190, 105408. <https://doi.org/10.1016/j.cmpb.2020.105408>

- Rosero, A. C. T., & Álvarez, C. C. B. (2012). Riesgos psicosociales intralaborales en docencia. *Revista Iberoamericana de Psicología: Ciencia y Tecnología.*, 5(2), 95–106.
- Rosset, S., & Tibshirani, R. J. (2020). From Fixed-X to Random-X Regression: Bias-Variance Decompositions, Covariance Penalties, and Prediction Error Estimation. *Journal of the American Statistical Association*, 115(529), 138–151.
<https://doi.org/10.1080/01621459.2018.1424632>
- Rumelhart, D. E., & MacClelland, J. L. (Eds). (1986). *Parallel Distributed Processing* (Vol. 1). MIT Press.
- Salal, Y. K., Abdullaev, S. M., & Kumar, M. (2019). Educational Data Mining: Student Performance Prediction in Academic. *International Journal of Engineering and Advanced Technology*, 8(4), 54–59.
- Santos, M. (2011). Un Enfoque Aplicado del Control Inteligente. *Revista Iberoamericana de Automática e Informática Industrial RIAI*, 8(4), 283–296.
<https://doi.org/10.1016/j.riai.2011.09.016>
- Sauter, S. L., & Murphy, L. R. (1984). Factores psicosociales y de organización. *Enciclopedia de salud y seguridad en el trabajo*, 34.2-34.75.
- Scholkopf, B., Smola, A. J., & Bach, F. (2018). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. MIT Press.
- Schweidtmann, A. M., & Mitsos, A. (2019). Deterministic Global Optimization with Artificial Neural Networks Embedded. *Journal of Optimization Theory and Applications*, 180(3), 925–948. <https://doi.org/10.1007/s10957-018-1396-0>
- Siegrist, J. (1996). Adverse Health Effects of High-Effort/Low-Reward Conditions. *Journal of Occupational Health Psychology*, 1(1), 27–41.

- Silva Gutiérrez, B. N., & Vicente Flores, R. (2015). Académicos universitarios y salud ocupacional. Importancia de los factores psicosociales y variables sociodemográficas, el caso de un centro universitario regional de la Universidad de Guadalajara, México. *Población y Desarrollo - Argonautas y Caminantes*, 10, 33–43. <https://doi.org/10.5377/pdac.v10i0.1736>
- Sriramprakash, S., Prasanna, V. D., & Murthy, O. V. R. (2017). Stress Detection in Working People. *Procedia Computer Science*, 115, 359–366. <https://doi.org/10.1016/j.procs.2017.09.090>
- Suárez Sánchez, A., Riesgo Fernández, P., Sánchez Lasheras, F., de Cos Juez, F. J., & García Nieto, P. J. (2011). Prediction of work-related accidents according to working conditions using support vector machines. *Applied Mathematics and Computation*, 218(7), 3539–3552. <https://doi.org/10.1016/j.amc.2011.08.100>
- Subhani, A. R., Mumtaz, W., Saad, M. N. B. M., Kamel, N., & Malik, A. S. (2017). Machine Learning Framework for the Detection of Mental Stress at Multiple Levels. *IEEE Access*, 5, 13545–13556. <https://doi.org/10.1109/ACCESS.2017.2723622>
- Sushant K, S. (2020). A Commentary on the Application of Artificial Intelligence in the Insurance Industry. *Trends in Artificial Intelligence*, 4(1). <https://doi.org/10.36959/643/305>
- Tettamanzi, A., & Tomassini, M. (2001). *Soft Computing*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-04335-6>
- Teuscher, C. (Ed.). (2004). *Alan Turing: Life and Legacy of a Great Thinker*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-05642-4>
- Tiffin, P. A., & Paton, L. W. (2018). Rise of the machines? Machine learning approaches and mental health: opportunities and challenges. *The British Journal of Psychiatry*, 213(3), 509–510. <https://doi.org/10.1192/bjp.2018.105>

- Tuttle, J. F., Vesel, R., Alagarsamy, S., Blackburn, L. D., & Powell, K. (2019). Sustainable NO_x emission reduction at a coal-fired power station through the use of online neural network modeling and particle swarm optimization. *Control Engineering Practice*, 93, 104167. <https://doi.org/10.1016/j.conengprac.2019.104167>
- Tyson, W. R., & Miller, W. A. (1977). Surface free energies of solid metals: Estimation from liquid surface tension measurements. *Surface Science*, 62(1), 267–276. [https://doi.org/10.1016/0039-6028\(77\)90442-3](https://doi.org/10.1016/0039-6028(77)90442-3)
- Tzeng, H.-M., Hsieh, J.-G., & Lin, Y.-L. (2004a). *Predicting Nurses' Intention to Quit With a Support Vector Machine*. 11.
- Tzeng, H.-M., Hsieh, J.-G., & Lin, Y.-L. (2004b). Predicting Nurses' Intention to Quit With a Support Vector Machine. *Computers, Informatics, Nursing*, 22(4), 232–242.
- Uronen, L., Moen, H., Teperi, S., Martimo, K.-P., Hartiala, J., & Salanterä, S. (2020). Towards automated detection of psychosocial risk factors with text mining. *Occupational Medicine*, 70(3), 203–206. <https://doi.org/10.1093/occmed/kqaa022>
- Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer New York. <https://doi.org/10.1007/978-1-4757-2440-0>
- Vieco Gómez, G. F., & Abello Llanos, R. (2014). Psychosocial factors at work, stress and morbidity around the world. *Psicología desde el Caribe*, 31(2), 354–385. <https://doi.org/10.14482/psdc.31.2.5544>
- Villalobos F, G. H. (2004). Vigilancia Epidemiológica de los Factores Psicosociales. Aproximación Conceptual y Valorativa. *Ciencia & Trabajo*, 6(14), 194–201.
- Villalobos, G. (2005). *Diseño de un sistema de vigilancia epidemiológica de factores de riesgo psicosocial en el trabajo*. Escuela Nacional de Salud Pública.
- Villalobos, G., Vargas M, A., Escobar, J., Jiménez, M. L., & Rondón, M. A. (2010). *Batería de instrumentos para la evaluación de factores de riesgo psicosocial* (1a ed., Vol.

- 1). Ministerio de protección social.
<http://fondoriesgoslaborales.gov.co/documents/publicaciones/estudios/>
- Wang, S., Cao, Y., Huang, T., Chen, Y., Li, P., & Wen, S. (2020). Sliding mode control of neural networks via continuous or periodic sampling event-triggering algorithm. *Neural Networks*, *121*, 140–147. <https://doi.org/10.1016/j.neunet.2019.09.001>
- Wang, X., Wang, Z., Song, Q., Shen, H., & Huang, X. (2020). A waiting-time-based event-triggered scheme for stabilization of complex-valued neural networks. *Neural Networks*, *121*, 329–338. <https://doi.org/10.1016/j.neunet.2019.09.032>
- Weissbrodt, R., & Giauque, D. (2017). Labour inspections and the prevention of psychosocial risks at work: A realist synthesis. *Safety Science*, *100*, 110–124. <https://doi.org/10.1016/j.ssci.2017.02.012>
- Wenhui Liao, Weihong Zhang, Zhiwei Zhu, & Qiang Ji. (2005). A Real-Time Human Stress Monitoring System Using Dynamic Bayesian Network. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, *3*, 70–70. <https://doi.org/10.1109/CVPR.2005.394>
- Werboz, P. J. (1974). *Beyond Regression: New Tools for Prediction and Analysis in Behavioral Sciences*. Harvard.
- Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits. *Stanford Univ Ca Stanford Electronics Labs*.
- Wikimedia Commons. (2006, agosto 26). *Waterstrider—Aquarius remigis*.
WaterstriderEnWiki.Jpg.
<https://commons.wikimedia.org/wiki/File:WaterstriderEnWiki.jpg>
- Yadav, S. K., & Hashmi, A. (2018). An Investigation of Occupational stress Classification by using Machine Learning Techniques. *International Journal of Computer Sciences and Engineering*, *6(6)*, 842–850.
<https://doi.org/10.26438/ijcse/v6i6.842850>

- Yang, X.-S. & Suash Deb. (2009). Cuckoo Search via Levy flights. *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)*, 210–214.
<https://doi.org/10.1109/NABIC.2009.5393690>
- Ye, L., Ferdinando, H., & Alasaarela, E. (2014). Techniques in Pattern Recognition for School Bullying Prevention: Review and Outlook. *Journal of Pattern Recognition Research*, 9(1), 50–63. <https://doi.org/10.13176/11.586>
- Yigit, I. O., & Shourabizadeh, H. (2017). An approach for predicting employee churn by using data mining. *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, 1–4. <https://doi.org/10.1109/IDAP.2017.8090324>
- Zhou, G., Moayedi, H., Bahiraei, M., & Lyu, Z. (2020). Employing artificial bee colony and particle swarm techniques for optimizing a neural network in prediction of heating and cooling loads of residential buildings. *Journal of Cleaner Production*, 254, 120082. <https://doi.org/10.1016/j.jclepro.2020.120082>
- Zhou, H., Wang, J., Wu, J., Zhang, L., Lei, P., & Chen, X. (2013). Application of the Hybrid SVM-KNN Model for Credit Scoring. *2013 Ninth International Conference on Computational Intelligence and Security*, 174–177.
<https://doi.org/10.1109/CIS.2013.43>
- Zubair, M., & Yoon, C. (2020). Multilevel mental stress detection using ultra-short pulse rate variability series. *Biomedical Signal Processing and Control*, 57, 101736.
<https://doi.org/10.1016/j.bspc.2019.101736>

VIII. 6. Anexo A: Proyecto de Investigación ejecutado

“Convocatoria Nacional para el Apoyo al Desarrollo de Tesis de Posgrado o de Trabajos Finales de Especialidades en el área de la Salud de la Universidad Nacional de Colombia 2017-2018” Resolución 21 de 2017. Oficina de la Vice-rectoría de Investigación (21 de diciembre de 2017), proyecto seleccionado, financiado y ejecutado con recursos de la Universidad Nacional de Colombia con el número de identificación 40976 en el Sistema de Información Hermes. Tiempo de ejecución 12 meses.

Vicerrectoría de Investigación



ACTA DE FINALIZACIÓN DE PROYECTO DE INVESTIGACIÓN

Sandra Ximena Carvajal Quintero, en su calidad de Vicedecana de Investigación y Extensión Facultad de Ingeniería y Arquitectura considerando que:

- A través de la CONVOCATORIA NACIONAL PARA EL APOYO AL DESARROLLO DE TESIS DE POSGRAD TRABAJOS FINALES DE ESPECIALIDADES EN EL ÁREA DE LA SALUD, DE LA UNIVERSIDAD NACIONAL DE COL. 2017-2018 en la modalidad: Modalidad 1: Apoyo a tesis de doctorado en cualquier área del conocimiento, se aprobó el proyecto Sistema de clasificación basado en técnicas inteligentes para identificar el grado de riesgo psicosocial docentes de educación básica primaria y secundaria en colegios públicos de Colombia, cuyo investigador principal es el (la) profesor(a) OMAR DANILO CASTRILLON GOMEZ.
- El proyecto se financió por un total de \$15,000,000.
- El investigador entregó los informes parciales y final previstos y demás productos derivados del proyecto de investigación.
- De las evaluaciones a los informes finales, técnico y financiero, elaboradas por la Dirección de Investigación de la sede, se deduce que el investigador principal cumplió satisfactoriamente con los compromisos estipulados en el Acta de Iniciación y su ejecución fue de \$15,000,000.
- El término de la duración del proyecto fue hasta el 10 de septiembre de 2019.

Ha acordado:

PRIMERO: Dar por finalizado el proyecto con código Hermes 40976 y código QUIPUJ 201010016754 por haberse cumplido los compromisos adquiridos satisfactoriamente.

SEGUNDO: Declarar al investigador principal a paz y salvo con el (la) FACULTAD DE INGENIERÍA Y ARQUITECTURA por todo concepto relacionado con el proyecto con código 40976.

En constancia se firma la presente acta en la ciudad de Bogotá el día 24 de septiembre de 2020

Cordialmente,

Sandra Ximena Carvajal Quintero
Vicedecana de Investigación y Extensión Facultad de Ingeniería y
Arquitectura

OMAR DANILO CASTRILLON GOMEZ
Director del proyecto

HER- 61 - 2020



sep 24, 2020 3:42 p.m.

Pág. 1 de 1

IX. Anexo B: Pasantía investigativa

Como parte del proceso de formación doctoral, se realizó una pasantía investigativa de 7 meses en la Universidad Politécnica de Valencia, Comunidad Valenciana, España durante el año 2018.



Valencia, 27 de noviembre de 2018

Apreciados
Universidad Nacional de Colombia,
Sede Manizales
Manizales, Colombia

Apreciados

Con la siguiente doy fe que el estudiante de doctorado Rodolfo Mosquera Navarro, de nacionalidad Colombiana, realizó satisfactoriamente una estancia de investigación en nuestro Centro de Gestión de la Calidad y del Cambio (CQ), perteneciente al Departamento de Estadística e Investigación Operativa Aplicadas y Calidad en la Universidad Politécnica de Valencia, dentro de las actividades de cooperación que desde hace años desarrollan nuestra dos universidades.

Durante su estancia le invitamos a participar como doctorante invitado dentro del grupo de investigación tomó cursos y realizó labor con nuestro grupo docente con excelente desempeño.

La visita tuvo lugar entre los días 25 de abril de 2018 y 27 de octubre de 2018, según su conveniencia.

Reciban un cordial saludo,



Andrés Carrión García
Director, Centro de gestión de la Calidad y del Cambio
Universidad Politécnica de Valencia
Comunidad Valenciana - España



Valencia, 27 de noviembre de 2018

X. Anexo C: Estrategia de búsqueda sistemática, ecuaciones de búsqueda, resultados obtenidos y distribución de publicaciones por año

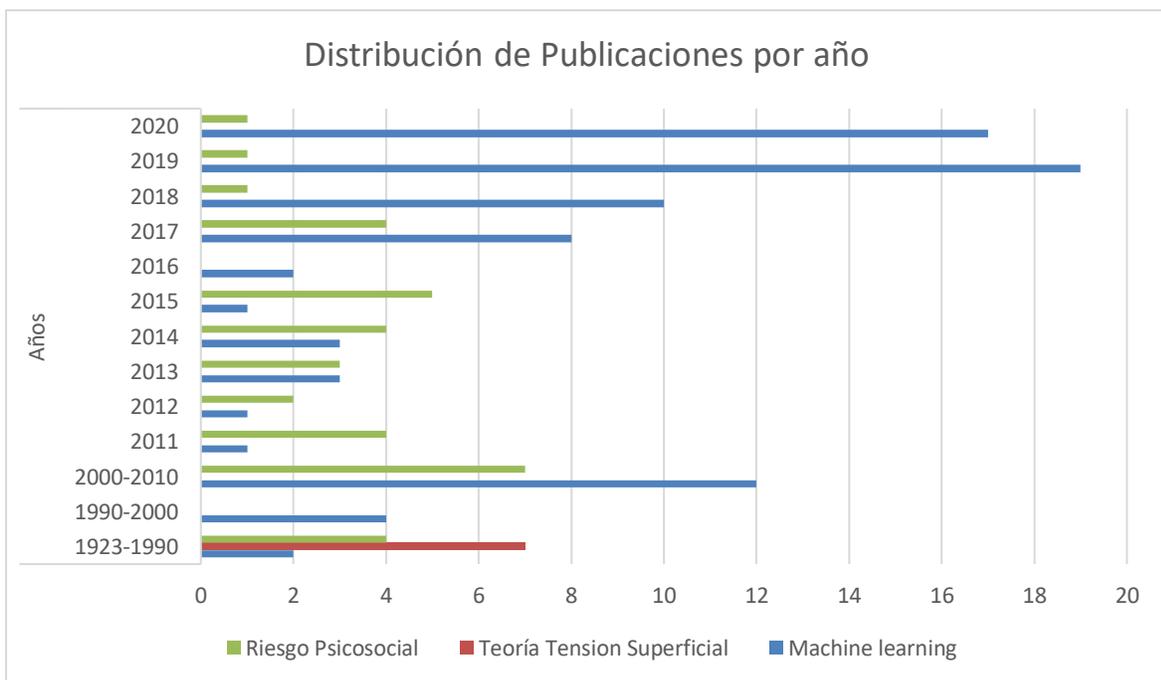
Desde el año 1990 hasta la fecha de aplicación.

Tema	Subtema (AND; OR)	Enfoque (AND; OR)
	<u>Contenido</u>	
“data mining”	“psychosocial risks”	Artificial intelligence or algorithms
	“Hazard identification”	
	“Risk*”	
	Identification and risk assessment	
	Evaluation*	
	Hazard identification and risk assessment and control	
“machine learning”	“psychosocial risks”	Artificial intelligence or algorithms
	“Hazard identification”	
	“Risk*”	
	Identification and risk assessment	
	Evaluation*	

	Hazard identification and risk assessment and control	
“psychosocial risks”	Algorithms	“Machine learning”
“psychosocial factors”		
“university staff”		
“teachers”		
“psychosocial risks”	“Hazard identification”	“Fuzzy sets”
“psychosocial factors”	“Risk*”	“Fuzzy systems”
“university staff”	Hazard identification and risk assessment and control	“Fuzzy logic with engineering applications”
“teachers”	Identification and risk assessment	“fuzzy theory”
“riesgo psicosocial”	Evaluation*	Neural network
“docentes”		
	<u>Proceso</u>	
	“Fuzzy risk modeling*”	“Neural network classifier”
	Fuzzy model for assessing risk of occupational safety	Bayesian network classifiers
	Fuzzy inference system	Bayesian classifier*
	Fuzzy logic and applications	“Bayesian approach”
	Fuzzy logic and neuro-fuzzy modeling	Bayesian networks
	Neural classifier*	
	Neural network classification	
	Bayesian analysis	
	Data minning	psychosocial risks
	Data minning	psychosocial factors
	Machine learning	Artificial neural networks
	Machine learning	psychosocial risks

	Machine learning	psychosocial factors
	Artificial neural networks	psychosocial risks
	Artificial neural networks	psychosocial factors
	Classification	psychosocial risks
	Classification	psychosocial factors
	Prediction	psychosocial risks
	Prediction	psychosocial factors
	Predict	psychosocial risks
	Predict	psychosocial factors
	Artificial neural networks	teachers
	Data minning	teachers
	Machine learning	teachers
	Prediction	teachers
	Predict	teachers

- Distribución de publicaciones por año



XI. Anexo D: Protocolo para la aplicación de la batería de riesgo psicosocial y para la toma de señales y formato de consentimiento informado

Secuencia de la aplicación:

Se siguió la siguiente secuencia de actividades:

- a. Presentación.
- b. Carta de información y firma del consentimiento informado.
- c. Aplicación del cuestionario.
- d. Aplicación de la ficha de datos generales.

Toma de muestra de señales:

- a. Informe al paciente sobre la puesta de electrodos en diferentes partes.
- b. Informe al paciente sobre que la toma de señales no será dolorosa ni incurrirá en el paso de corriente a través de la piel.

c. Informe al paciente que las señales aquí obtenidas será para uso exclusivo de la investigación.

Docentes pertenecientes a diferentes áreas y diferentes instituciones que colaboraron de forma voluntaria en la investigación.



	CONSENTIMIENTO INFORMADO DE LA APLICACIÓN DE LA BATERIA DE RIESGO PSICOSOCIAL PROYECTO DE INVESTIGACIÓN	FECHA DE DILIGENCIAMIENTO
		AAAA - MM - DD

Yo, _____ servidor público identificado(a) con cédula de ciudadanía No. _____ de _____, declaro haber sido plenamente informado (a) sobre el proceso de evaluación de los Factores de Riesgo Psicosocial que adelantará el **COLEGIO XXXXXXXXXX**, en cumplimiento de la Resolución 2646 de 2008, el cual consta de lo siguiente:

1. La evaluación tiene como objetivo la identificación y valoración de los factores psicosociales y sus posibles efectos sobre mi salud y/o desempeño, con el fin de establecer planes de Intervención que respondan a las características y necesidades detectadas en los servidores públicos de la entidad.
2. La metodología de evaluación consiste en la aplicación de 1 instrumento de medición de los factores de riesgo psicosocial, debidamente validados para la población trabajadora colombiana, dichos instrumentos son:
 - ✓ Ficha de datos generales
 - ✓ Cuestionario de factores de riesgo psicosocial intralaboral
 - ✓ Medición de problemas osteomusculares a través de dispositivo para electromiografía.
 - ✓ Medición de la respuesta galvanica de la piel para observar el comportamiento positivo o negativo a la respuesta.
3. Los resultados obtenidos serán de absoluta RESERVA. Y solo se utilizarán para fines científicos de la investigación.
4. La información individual utilizada para la presente evaluación será manejada con absoluta confidencialidad y será usada exclusivamente para los fines inherentes a la Seguridad y Salud en el Trabajo. La misma, será revelada únicamente con mi consentimiento o el de mi representante legal, excepto en aquellas circunstancias particulares en que no hacerlo represente un evidente daño para mí o para otros.
5. Mi participación en esta evaluación es estrictamente voluntaria; es decir que yo puedo libremente decidir participar en el proceso o rehusarme a hacerlo, sin que esta decisión represente ningún perjuicio para mi persona.
6. Me han informado que puedo formular las preguntas sobre el proceso de evaluación, en cualquier momento del mismo, al profesional que acompaña el proceso (Psicólogo con posgrado en Seguridad y Salud en Trabajo).
7. En el evento en que el profesional a cargo lo requiera, podrá contactarme por vía telefónica o correo electrónico.

Con base en la información que me han suministrado sobre el proceso de evaluación de los factores psicosociales, declaro libremente que he leído y entendido completamente lo referente al mismo (objetivos, metodología, tratamiento de la información, informes de resultados, participación voluntaria). De igual manera declaro que conozco mi derecho a formular las preguntas que me surjan antes, durante y/o después del proceso de evaluación, con el fin de aclarar lo pertinente.

Teniendo en cuenta lo anterior, expreso libremente mi voluntad de:

(Marque una X en el recuadro que corresponda, según su decisión y a continuación registre sus datos)

SI Participar en el proceso de evaluación de factores de Riesgo Psicosocial

NO Participar en el proceso de evaluación de factores de Riesgo Psicosocial

Nombre del Servidor Público: _____

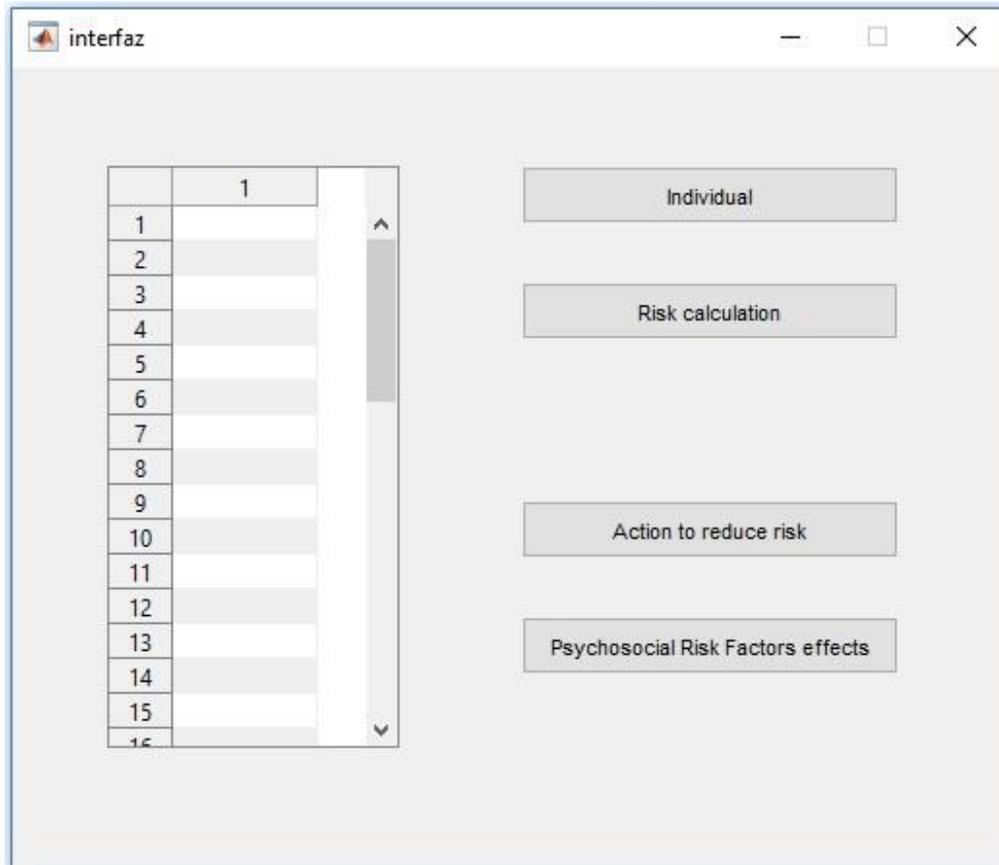
Documento de Identificación: _____

Correo electrónico: _____

FIRMA

XII. Anexo E: Interfaz para la gestión del riesgo psicosocial basado en el algoritmo de Red neuronal de tensión superficial

El disponer de una interfaz facilita que el psicólogo laboral pueda elaborar las evaluaciones de una forma comoda y obtener el resultado de predicción basado en el algoritmo.



XIII. Anexo F: Código algoritmo de Ascenso de colinas con reinicio aleatorio – Máquinas de soporte vectorial (RRHC-SVM)

```

% Lectura de datos del excel
%
-----

disp('--- Espere un momento por favor ---')
disp('--- Procesando ---')
h=msgbox('procesando datos...');
entrada=xlsread('Datos experimentales.xlsx','Datos experimentales1');
columnadz=entrada(:,end);
entrada(:,end)=[];
load borrar5 % borrar5 tiene entrada y columnadz

%
-----

% Selección del 60% de los datos como training set
%
-----

porcentaje = 0.6;
% trainingset =
sort([randsample(find(columnadz==1),ceil(sum(columnadz==1)*porcentaje));...
%      randsample(find(columnadz==2),ceil(sum(columnadz==2)*porcentaje));...
%      randsample(find(columnadz==3),ceil(sum(columnadz==3)*porcentaje));...
%      randsample(find(columnadz==4),ceil(sum(columnadz==4)*porcentaje))]);
load borrar6 % borrar6 tiene trainigset (esto es para trabajar siempre con los mismos:
tamaño 290)
validationset = setdiff(1:size(entrada,1),trainingset)';

entradatr = entrada(trainingset,:); % Los valores de entrada correspondientes al training
set
columnadztr = columnadz(trainingset); % Lo valores de columnadz correspondientes al
training set

entradavl = entrada(validationset,:); % Los valores de entrada correspondientes al
training set
columnadzvl = columnadz(validationset); % Lo valores de columnadz correspondientes al
training set

%
-----

```

```
% Filtrado de variables redundantes
%
```

```
objetos = entradatr'; % En entrada están las 129 variables independientes ubicadas en
filas
Dis = pdist(objetos,'chebychev'); % Métrica de similitud usada: distancia chebychev o
métrica máxima
Z = linkage(Dis,'average'); % El criterio de enlace promedio busca los pares de variables
que se encuentran más cercanas entre sí dependiendo de lo obtenido en Y
T = cluster(Z,'maxclus',30); % Se determina dónde cortar el árbol de jerarquía para crear
los clusters. En esta ocasión, se busca crear 30 clusters.
```

```
%
```

```
% Selección de las variables representativas de cada cluster
%
```

```
eT = unique(T); % Se crea un vector con los valores del 1 al 30, que son los clusters
creados
grupos = cell(1,length(eT)); % Se genera una variable inicial para guardar las variables
correspondientes a cada cluster
gr = zeros(1,length(eT)); % Se genera una variable inicial para guardar las variables
representantes de cada cluster
for i=1:length(eT) % Se genera un cell array donde se guardan las variables de cada
cluster
    grupos{i} = find(eT(i)==T);
end
for i=1:length(grupos) % Se va a verificar la operación por cada cluster
    if length(grupos{i})~=1 % Si en el cluster hay más de una variable
        a=pdist(entrada(:,grupos{i}));
        aa=squareform(a);
        aaa=sum(aa);
        [~,elques]=min(aaa);
        gr(i) = grupos{i}(elques);
    elseif length(grupos{i})==1 % Si en el cluster hay 1 variable
        gr(i) = grupos{i}(1);
    end
end
end
gr=sort(gr); % En gr quedan las variables representativas de cada cluster
```

```
%
```

```
% Análisis de componentes principales
%
```

```
[coeff,nuevosdatos,~,~,explained,mu] = pca(entradatr(:,gr));
```

```
%
```

```
% Selección de componentes principales
```

```

%
-----

primeracol = nuevosdatos(:,1);
segundacol = nuevosdatos(:,2);
% Graficación (opcional)
%scatter(primeracol,segundacol,25,columnadztr,'filled') % columnadz corresponde a la
variable independiente

% Etiquetado de datos con Riesgo=1 y Riesgo=2, con 1 y -1 respectivamente
grupor1 = [primeracol(columnadztr==1) segundacol(columnadztr==1)]; % 1c: primeracol,
2c: segundacol
grupor2 = [primeracol(columnadztr==2) segundacol(columnadztr==2)]; % 1c: primeracol,
2c: segundacol
grupor3 = [primeracol(columnadztr==3) segundacol(columnadztr==3)]; % 1c: primeracol,
2c: segundacol
grupor4 = [primeracol(columnadztr==4) segundacol(columnadztr==4)]; % 1c: primeracol,
2c: segundacol

% Grupos de entrenamiento de clasificadores
gtraining = [1 2
             2 3
             3 4];

% Etiquetas
eti = [1 -1];

separadores=cell(1,size(gtraining,1)); % Se inicializa una variable para los hiperplanos
separadores

X = {};
Y = {};

for i=1:size(gtraining,1)
    % Concatenación de los grupos con los cuales se va a trabajar
    % En X se guardan los datos obtenidos de pca
    % En Y se guardan las etiquetas
    eval(['X{i} = [ones(size(grupor' num2str(gtraining(i,1)) ',1)+size(grupor'
num2str(gtraining(i,2)))...
',1),1) vertcat(grupor' num2str(gtraining(i,1)) ',grupor' num2str(gtraining(i,2)) ');'];]);
    eval(['Y{i} = vertcat(repmat(eti(1),size(grupor' num2str(gtraining(i,1)))...
',1),1),repmat(eti(2),size(grupor' num2str(gtraining(i,2)) ',1),1));'];]);

%
-----

% Ascenso de colina con reinicio aleatorio
%
-----

% Generación de un individuo inicial al azar y cálculo de su error
% individuo = [1 1 1];
% error = malclasificados(individuo,X{i},Y{i});

```

```

% Este if se utiliza en caso de que ya exista en valor de individuo
% guardado en una base de datos, si ya existe se utiliza ese valor como
% individuo inicial al azar, si no existe el individuo inicial al azar
% será individuo=[1 1 1];
if ~exist('actualizacion_individuo','file')
    individuo=[1 1 1];
    error = malclasificados(individuo,X{i},Y{i});
    save('actualizacion_individuo','individuo','error')
else
    load 'actualizacion_individuo'
end

for kglobal=1:1000000
    individuora = -10000 + (20000)*rand(1,3);
    errorra = malclasificados(individuora,X{i},Y{i});
    if errorra<error
        individuo = individuora;
        error = errorra;
        save('actualizacion_individuo','individuo','error')
        disp(['encontré uno mejor con error' num2str(error)])
        if error == 0
            disp('break1'), break
        end
    end
%%% Inicio - Ascenso de colina
    for klocal=1:10000
        individuo2 = individuo;
        individuo2(randi(3)) = -10000 + (20000)*rand;
        error2 = malclasificados(individuo2,X{i},Y{i});
        if error2<error
            individuo = individuo2;
            error = error2;
            save('actualizacion_individuo','individuo','error')
            disp(['muté uno mejor con error:' num2str(error)])
            if error == 0
                disp('break2'), break
            end
        end
    end
end
%%% Fin - Ascenso de colina

end
end

separadores{i} = individuo;

end

graficar = 1;
figure
if graficar == 1

```

```

hold on
plot(grupor1(:,1),grupor1(:,2),'bx')
hold on, plot(grupor2(:,1),grupor2(:,2),'mx')
hold on, plot(grupor3(:,1),grupor3(:,2),'gx')
hold on, plot(grupor4(:,1),grupor4(:,2),'rx')

for ff=1:length(separadores)
    hold on, plot([-200 250], [(-separadores{ff}(2)/separadores{ff}(3))*-200 + (-
separadores{ff}(1)/separadores{ff}(3)) (-separadores{ff}(2)/separadores{ff}(3))*250 + (-
separadores{ff}(1)/separadores{ff}(3))], '--')
    end

    axis([-200 250 -20 30])
accuracy=0;
for i=1:size(gtraining,1)
    eval(['X = [ones(size(grupor' num2str(gtraining(i,1)) ' ,1)+size(grupor'
num2str(gtraining(i,2))...
',1),1) vertcat(grupor' num2str(gtraining(i,1)) ',grupor' num2str(gtraining(i,2)) ');'];]);
    eval(['Y = vertcat(repmat(etiq(1),size(grupor' num2str(gtraining(i,1))...
',1),1),repmat(etiq(2),size(grupor' num2str(gtraining(i,2)) ',1),1));'];]);
    hold on,
quack=fitcsvm(X(:,2:3),Y,'Standardize',true,'KernelFunction','Polynomial','KernelScale','auto');
    hold off
    figure
accuracy=accuracy+(sum(Y==predict(quack,X(:,2:3)))/length(Y));
    plotconfusion(categorical(Y),categorical(predict(quack,X(:,2:3))))

    title(['Clasificación Separador ' num2str(i)])
end
accuracy=accuracy/size(gtraining,1);
fprintf('Accuracy = %f%%\n',accuracy*100)

end

```

XIV. Anexo G: Código algoritmo de Búsqueda cuckoo – Máquinas de soporte vectorial (CS-SVM)

% ALGORITMO DE CLASIFICACIÓN CON BÚSQUEDA CUCKOO

clearvars
%

% Entrada de datos
%

load datos0c % Tiene entrada y CLASE
disp('--- Espere un momento por favor ---')
disp('---Cuckoo Search-SVM---')
disp('--- Procesando ---')
h=msgbox('procesando datos...');
%

% Selección del 60% de los datos como training set
%

porcentaje = 0.6;
trainingset = sort([randsample(find(CLASE==1),ceil(sum(CLASE==1)*porcentaje));...
 randsample(find(CLASE==2),ceil(sum(CLASE==2)*porcentaje));...
 randsample(find(CLASE==3),ceil(sum(CLASE==3)*porcentaje));...
 randsample(find(CLASE==4),ceil(sum(CLASE==4)*porcentaje))]);

validationset = setdiff(1:size(entrada,1),trainingset)';

entradatr = entrada(trainingset,:); % Los valores de entrada correspondientes al training set
CLASEtr = CLASE(trainingset); % Lo valores de columnadz correspondientes al training set

entrada vl = entrada(validationset,:); % Los valores de entrada correspondientes al training set
CLASE vl = CLASE(validationset); % Lo valores de columnadz correspondientes al training set

%

% Preparación de los datos para el algoritmo de búsqueda

```

%
-----

grupor1 = entradatr(CLASEtr==1,:);
grupor2 = entradatr(CLASEtr==2,:);
grupor3 = entradatr(CLASEtr==3,:);
grupor4 = entradatr(CLASEtr==4,:);

% Grupos de entrenamiento de clasificadores
gtraining = [1 2
             2 3
             3 4];

% Etiquetas
etiq = [1 -1];

separadores=cell(1,size(gtraining,1)); % Se inicializa una variable para los hiperplanos
separadores

X = {};
Y = {};

%
-----

% Cuckoo search
%
-----

% Parámetros de cuckoo search
nidos = 5; % Número de nidos
pa = 0.1; % Porcentaje de peores
generaciones = 2000;

for i=1:size(gtraining,1)

    % Concatenación de los grupos con los cuales se va a trabajar
    % En X se guardan los datos
    % En Y se guardan las etiquetas
    eval(['X{i} = [ones(size(grupor' num2str(gtraining(i,1)) ',1)+size(grupor'
num2str(gtraining(i,2)))...
',1),1) vertcat(grupor' num2str(gtraining(i,1)) ',grupor' num2str(gtraining(i,2)) ');'];
    eval(['Y{i} = vertcat(repmat(etiq(1),size(grupor' num2str(gtraining(i,1)))...
',1),1),repmat(etiq(2),size(grupor' num2str(gtraining(i,2)) ',1),1));'];

% Se calculan individuos al azar, y sus márgenes
pobcuckoo = zeros(nidos,size(X{i},2));
marcuckoo = zeros(nidos,1);
for c = 1:nidos
    pobcuckoo(c,:) = perceptron(X{i}(:,2:end),Y{i}); % Individuos
    marcuckoo(c,1) = margen(pobcuckoo(c,:),X{i}(:,2:end)); % Márgenes
end

```

```

for it = 1:generaciones

%   Se seleccionan dos individuos al azar (posiciones al azar)
pos = randsample(1:nidos,2);
% i -> pos(1)
% j -> pos(2)

%   Se crea una matriz con los nidos i, j
nidoi = pobcuckoo(pos(1),:);
nidoj = pobcuckoo(pos(2),:);

%   Se escoge la posición del huevo intruso en el nido i
huevointruso = randi(size(pobcuckoo,2));

%   Se hace la modificación de Levy Flight al huevo intruso,
%   se asegura que el cambio genere un buen separador (error 0)
error = 1;
while error ~= 0
    nidoi(huevointruso) = levyf(pobcuckoo(pos(1),huevointruso));
    error = malclasificados(nidoi,X{i},Y{i});
end

%   Se calculan los márgenes de ambos nidos
margini = margen(nidoi,X{i}(:,2:end));
marginj = margen(nidoj,X{j}(:,2:end));

%   Condicional, si (Margen de i) > (Margen de j)
%   se reemplaza el nido j con el nido i
if margini > marginj
%   Se hace el reemplazo del mejor individuo con su margen
pobcuckoo(pos(2),:) = nidoi;
marcuckoo(pos(2),1) = margini;
end

%   Se saca un porcentaje de los peores, y se reemplazan con soluciones
%   nuevas al azar
[~,idx2] = sort(marcuckoo);
for peores = 1:ceil(nidos*pa)
    pobcuckoo(idx2(peores),:) = perceptron(X{i}(:,2:end),Y{i});
    marcuckoo(idx2(peores),1) = margen(pobcuckoo(idx2(peores),:),X{i}(:,2:end));
end

%   Se calcula el mejor cuckoo de la generación actual
[margenmaximo{i}(it),posbestcuckoo] = max(marcuckoo);
bestcuckoo{i}(it,:) = pobcuckoo(posbestcuckoo,:); % El mejor cuckoo

%   i,it,marcuckoo % Para ver el progreso de las operaciones

end % End de cuckoo search

separadores{i} = bestcuckoo{i}(end,:);

```

```

Mseparadores(i) = margenmaximo{i}(end);

end

save('separadorescuckoo','separadores','Mseparadores')
save('iteracionescuckoo','bestcuckoo','margenmaximo')
save('datosXY','X','Y')

%

% Validación de clasificadores
%
-----

% Se calcula el score de componentes principales para los datos de
% validación y se evalúan los separadores
for v=1:size(entradavl,1)

    % Se evalúan los separadores
    if dot(separadores{1},[1 entradavl(v,:)]) > 0 % >
        % 'Persona es de riesgo 1'
        clasevl(v,1) = 1;
    elseif dot(separadores{2},[1 entradavl(v,:)]) > 0 % >
        % 'Persona es de riesgo 2'
        clasevl(v,1) = 2;
    elseif dot(separadores{3},[1 entradavl(v,:)]) > 0 % >
        % 'Persona es de riesgo 3'
        clasevl(v,1) = 3;
    elseif dot(separadores{2},[1 entradavl(v,:)]) < 0 % <
        % 'Persona es de riesgo 4'
        clasevl(v,1) = 4;
    end
end
end
b=figure;
plotconfusion(categorical(CLASEvl),categorical(clasevl))
xlabel('Riesgo Verdadero')
ylabel('Riesgo Calculado')
title('Matriz de confusión')
tru=CLASEvl == clasevl;
eficiencia = 100*sum(tru)/length(CLASEvl);
disp('eficiencia')
disp(mean(eficiencia))
fprintf('%i bien clasificados de un total de %i\n',sum(tru),length(tru))
save('dbinterfaz','entrada','CLASE','separadores') % Base de datos para la interfaz

```

XV. Anexo H: Código algoritmo Red Neuronal de Tension Superficial (RNA-TS)

El código completo desarrollado en la versión de Matlab se puede consultar aquí: <https://codeocean.com/capsule/6532855/tree/v1> (Mosquera et al., 2019a). La versión 9.4 de Matlab fue usada para el desarrollo del algoritmo.

```
function procesoinicial
%la siguiente linea es la semilla guardada de los resultados
RandStream.setGlobalStream(RandStream('mt19937ar','Seed',387));
%si se quiere obtener nuevos resultados diferentes comentar la linea 3
blancoynegro=false;%si se coloca true las graficas salen en blanco y negro
if ~exist('basededatos.mat','file')%si no existe la base de datos en Matlab entonces se crea
[datos,datosnum]=xlsread('Datos Experimentales NA2.xlsx');%se lee el archivo de Excel
save('basededatos','datos','datosnum');%se guarda la base de datos de Excel en Matlab
else%si ya existe la base de datos en Matlab entonces solo se carga
%h=msgbox('procesando datos...');
h=msgbox({'Physical Surface Tension-Neural Net is Processing'; 'Please wait a moment
while all images load...'});

load basededatos %se carga la base de datos de Matlab
end
num=datos(:,1:end-1);%datos de entrada
try
nnet.guis.closeAllViews()%cerrada todas las vistas
close all%cerrar figura
clc%limpiar consola
catch
disp('algo inesperado sucedió al cerrar las vistas')
end
tar=zeros(4,size(num,2));%target
for k=1:size(num,2)
tar(datos(k,end),k)=1;%configurar target
end
net = patternnet([80 80 2]);%configuración inicial
net.performFcn='mse';%error cuadratico medio
net.TrainParam.showWindow=false;
net.TrainParam.max_fail=5000;%
net.TrainParam.epochs=5000;
net.divideparam.trainRatio=0.72;%datos de entrenamiento
net.divideparam.valRatio=0.08;%datos de validacion
```

```

net.divideparam.testRatio=0.2; %datos de evaluacion
net.biasConnect=[true; true; true; false]; %bias conectadas
net.layers{3}.transferFcn='purelin'; %funcion de transferencia
[net,tr] = train(net,num,tar);%entrenamiento y modificacion de arquitectura
[y,tt]=arquitectura(num,net);%agregado de capa de tension
coeffi=[22.1/1000 0.0477 0.0728 0.425]; %etanol, etilenglicol, agua, mercurio, tension
superficial
paraa=0.90;%temp
VH=y.*(coeffi(2)*paraa-coeffi(1)      coeffi(3)*paraa-coeffi(2)      coeffi(4)*paraa-coeffi(3)
0.34]'+coeffi';
[~,nivel]=max(y);%niveles
VH=VH(sub2ind(size(VH),nivel,1:size(VH,2)));
tt(:,2)=tt(:,1)./VH';
figure;
aa=plotperform(tr);
set(aa,'Name','Performance (Desempeño)')
title(' Iterations performance')
ylabel('Mean Square Error')
xlabel([num2str(tr.num_epochs) ' Iterations'])
legend({'Training','Validation','Test','Mejor'})
ss=findobj('DisplayName','Entrenamiento');
FF=get(ss,'XData');
FF2=get(ss,'YData');
set(ss,'XData',FF(1:100:end),'YData',FF2(1:100:end))
set(ss,'LineStyle','-','Marker','O','MarkerSize',5,'LineWidth',3)
if blancoynegro
    set(ss,'Color',[0 0 0])
end
ss=findobj('DisplayName','Validación');
set(ss,'LineStyle','-','Marker','+','MarkerSize',8,'LineWidth',2)
FF=get(ss,'XData');
FF2=get(ss,'YData');
set(ss,'XData',FF(1:100:end),'YData',FF2(1:100:end))
if blancoynegro
    set(ss,'Color',[0 0 0])
end
ss=findobj('DisplayName','Test');
if blancoynegro
    set(ss,'Color',[0 0 0])
end
FF=get(ss,'XData');
FF2=get(ss,'YData');
set(ss,'XData',FF(1:100:end),'YData',FF2(1:100:end))
set(ss,'LineStyle','-','Marker','.', 'MarkerSize',3,'LineWidth',3)
ss=findobj('DisplayName','Mejor');
delete(ss)
ss=findobj('MarkerSize', 16);
delete(ss)
figure
plotconfusion(tar,net(num))

```

```

title('Confusion Matrix')
xlabel('Target Class','FontWeight','bold','FontSize',12)
ylabel('Output Class','FontWeight','bold','FontSize',12);
xticklabels({'Risk 1','Risk 2','Risk 3','Risk 4',''})
yticklabels({'Risk 1','Risk 2','Risk 2','Risk 4',''})

FFF2=figure;
plotroc(tar,y)
title('Roc Curve','FontWeight','bold','FontSize',16)
xlabel('False Positives','FontWeight','bold','FontSize',12)
ylabel('True Positives','FontWeight','bold','FontSize',12)
set(legend,'String',strep(get(legend,'String'),'Class','Risk'))
ss=findobj(FFF2,'DisplayName','Risk 1');
if blancoynegro
    set(ss,'Color',[0 0 0])
end
FF=get(ss,'XData');
FF2=get(ss,'YData');
set(ss,'XData',FF(1:100:end),'YData',FF2(1:100:end))
set(ss,'LineStyle','-','Marker','O','MarkerSize',3,'LineWidth',3)
ss=findobj(FFF2,'DisplayName','Risk 2');
if blancoynegro
    set(ss,'Color',[0 0 0])
end
FF=get(ss,'XData');
FF2=get(ss,'YData');
set(ss,'XData',FF(1:100:end),'YData',FF2(1:100:end))
set(ss,'LineStyle','-','Marker','^','MarkerSize',5,'LineWidth',3)
ss=findobj(FFF2,'DisplayName','Risk 3');
if blancoynegro
    set(ss,'Color',[0 0 0])
end
FF=get(ss,'XData');
FF2=get(ss,'YData');
set(ss,'XData',FF(1:100:end),'YData',FF2(1:100:end))
set(ss,'LineStyle','-','Marker','d','MarkerSize',5,'LineWidth',3)
set(ss,'LineStyle','-','Marker','p','MarkerSize',5,'LineWidth',3)
ss=findobj(FFF2,'DisplayName','Risk 4');
if blancoynegro
    set(ss,'Color',[0 0 0])
end
FF=get(ss,'XData');
FF2=get(ss,'YData');
set(ss,'XData',FF(1:100:end),'YData',FF2(1:100:end))
set(ss,'LineStyle','-','Marker','.', 'MarkerSize',5,'LineWidth',3)

figure
% nivel=datos(:,end);
if blancoynegro==true
plot(tt(nivel==1,1),tt(nivel==1,2),'*k')
hold on

```

```

plot(tt(nivel==2,1),tt(nivel==2,2),'ok')
plot(tt(nivel==3,1),tt(nivel==3,2),'dk')
plot(tt(nivel==4,1),tt(nivel==4,2),'+k')
hold off
else
plot(tt(nivel==1,1),tt(nivel==1,2),'*b')
hold on
plot(tt(nivel==2,1),tt(nivel==2,2),'om')
plot(tt(nivel==3,1),tt(nivel==3,2),'dg')
plot(tt(nivel==4,1),tt(nivel==4,2),'+r')
hold off
end
title('Mass Vs Perimeter')
xlabel('Mass','FontWeight','bold','FontSize',12)
ylabel('Perimeter','FontWeight','bold','FontSize',12);
legend({'Risk 1','Risk 2','Risk 3','Risk 4'})
figure
if blancoynegro==true
kk=nthroot(tt(nivel==1,1)./tt(nivel==1,2),2);
plot(kk,1:length(kk),'ok','MarkerSize',5)
hold on
kk=nthroot(tt(nivel==2,1)./tt(nivel==2,2),2);
plot(kk,1:length(kk),'s','MarkerSize',3,'MarkerEdgeColor','k')
kk=nthroot(tt(nivel==3,1)./tt(nivel==3,2),2);
plot(kk,1:length(kk),'*k')
kk=nthroot(tt(nivel==4,1)./tt(nivel==4,2),2);
plot(kk,1:length(kk),'+k')
hold off
else
kk=nthroot(tt(nivel==1,1)./tt(nivel==1,2),2);
plot(kk,1:length(kk),'ob','MarkerSize',5)
hold on
kk=nthroot(tt(nivel==2,1)./tt(nivel==2,2),2);
plot(kk,1:length(kk),'s','MarkerSize',3,'MarkerEdgeColor','m')
kk=nthroot(tt(nivel==3,1)./tt(nivel==3,2),2);
plot(kk,1:length(kk),'*g')
kk=nthroot(tt(nivel==4,1)./tt(nivel==4,2),2);
plot(kk,1:length(kk),'+r')
hold off
end
title('$$${\boldmath$\sqrt{\text{mass/perimeter}}}$','Interpreter','latex','FontSize',16,'FontWeight','bold')
xlabel('$$${\boldmath$\sqrt{\text{mass/perimeter}}}$','Interpreter','latex','FontWeight','bold','FontSize',12)
ylabel('Data','FontWeight','bold','FontSize',12);
legend({'Risk 1','Risk 2','Risk 3','Risk 4','Location','northwest','FontWeight','bold'})
target=datos(:,end);
for k=1:4
    TP=sum(nivel(target==k)==k);
    FN=sum(nivel(target==k)~=k);

```

```

    FP=sum(nivel(target~=k)==k);
    TN=sum(nivel(nivel~=k)~=k);
    sensibilidad(k)=TP/(TP+FN);%true positive rate
    especificidad(k)=TN/(TN+FP);%true negative rate
    precision(k)=TP/(TP+FP);%positive predictive value
    prediconnegativa(k)=TN/(TN+FN);%negative predictive value
    errorclasificacion(k)=(FP+FN)/(FP+FN+TP+TN);%classification error
    accuracy(k)=(TP+TN)/(TP+TN+FP+FN);%accuracy
    prevalencia(k)=(TP+FN)/(TP+TN+FP+FN);%prevalence
end
disp('Sensitivity')
disp(mean(sensibilidad))
disp('Specificity')
disp(mean(especificidad))
disp('Precision')
disp(mean(precision(k)))
disp('Negative predictive value')
disp(mean(prediconnegativa))
disp('Classification Error')
disp(mean(errorclasificacion))
disp('Accuracy')
disp(mean(accuracy))
disp('Prevalence')
disp(mean(prevalencia))
end
function [Y,tt]=arquitectura(num,net)
X=num;%entrada
%parámetros de normalización;
x1_step1.xoffset =net.inputs{1}.processSettings{1}.xoffset;
x1_step1.gain = net.inputs{1}.processSettings{1}.gain;
x1_step1.ymin = net.inputs{1}.processSettings{1}.ymin;
b1 = net.b{1};
IW1_1 = net.IW{1};
b2 = net.b{2};
LW2_1 = net.LW{2};
b3 = net.b{3};
LW2_2 = net.LW{7};
LW2_3 = net.LW{12};
isCellX = iscell(X);
if ~isCellX
    X = {X};
end
TS = size(X,2); % timesteps
if ~isempty(X)
    Q = size(X{1},2); % samples/series
else
    Q = 0;
end
Y = cell(1,TS);
%calculo de las capas de salida
for ts=1:TS

```

```

Xp1 = mapminmax_apply(X{1,ts},x1_step1);
a1 = tansig_apply(repmat(b1,1,Q) + IW1_1*Xp1);
a2=tansig_apply(repmat(b2,1,Q) + LW2_1*a1);
a3=purelin_apply(repmat(b3,1,Q)+LW2_2*a2);
tt=a3'-min(a3')+1;
a4 = softmax_apply(LW2_3*a3);
Y{1,ts} = a4;
end
Xf = cell(1,0);
Af = cell(2,0);
if ~isCellX
    Y = cell2mat(Y);
end
end
function y = mapminmax_apply(x,settings)
y = bsxfun(@minus,x,settings.xoffset);
y = bsxfun(@times,y,settings.gain);
y = bsxfun(@plus,y,settings.ymin);
end
function a = softmax_apply(n,~)
if isa(n,'gpuArray')
    a = iSoftmaxApplyGPU(n);
else
    a = iSoftmaxApplyCPU(n);
end
end
function a = iSoftmaxApplyCPU(n)
nmax = max(n,[],1);
n = bsxfun(@minus,n,nmax);
numerator = exp(n);
denominator = sum(numerator,1);
denominator(denominator == 0) = 1;
a = bsxfun(@rdivide,numerator,denominator);
end
function a = iSoftmaxApplyGPU(n)
nmax = max(n,[],1);
numerator = arrayfun(@iSoftmaxApplyGPUHelper1,n,nmax);
denominator = sum(numerator,1);
a = arrayfun(@iSoftmaxApplyGPUHelper2,numerator,denominator);
end
function numerator = iSoftmaxApplyGPUHelper1(n,nmax)
numerator = exp(n - nmax);
end
function a = iSoftmaxApplyGPUHelper2(numerator,denominator)
if (denominator == 0)
    a = numerator;
else
    a = numerator ./ denominator;
end
end

```

```
function a = tansig_apply(n,~)
a = 2 ./ (1 + exp(-2*n)) - 1;
end
function a=purelin_apply(n,~)
a=n;
end
```

XVI. Anexo I: Carta de certificación por parte del psicólogo laboral que colaboró en el proyecto en la toma de datos

Santiago de Cali, octubre 12 de 2018

Asunto: Informe de aplicación de instrumento para evaluar riesgos psicosocial y de aplicación de software como herramienta de ayuda para la identificación de riesgo psicosocial.

Señor
RODOLFO MOSQUERA NAVARRO
CC 94527819
Ciudad

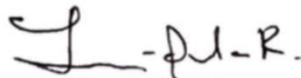
La siguiente es con el fin de informar a usted que se terminó de realizar la aplicación del instrumento al personal docente en donde se consideraron las variables de: **demandas del trabajo, control, liderazgo y relaciones sociales en el trabajo y recompensa**. La información recogida acorde a lo indicado en el código deontológico del psicólogo, se mantendrá confidencial bajo mi custodia.

También, según lo solicitado por usted y de acuerdo a la capacitación dada para el manejo del software, se aplicó la evaluación con los parámetros dados en el software y se comparó con los resultados anteriormente obtenidos en uno de los colegios. La información obtenida se manejó de manera anónima y se compararon los resultados por éste investigador, de lo cual puedo decir:

Los resultados dieron muy similares a la evaluación inicial realizada, el tiempo de ejecución del instrumento disminuyó substancialmente y se considera que la herramienta desarrollada contribuye para que los psicólogos organizacionales la utilicen. Se considera que la herramienta contribuye a determinar qué aspectos son los más relevantes frente a la percepción de los trabajadores para que la dirección pueda tomar decisiones sobre planes de mejoramiento de las condiciones del trabajo y prevención.

Agradezco la participación en el proyecto denominado **"Sistema de clasificación basado en técnicas inteligentes para identificar el grado de riesgo psicosocial en docentes de colegios públicos, área metropolitana de una ciudad colombiana"**

Atento a sus comentarios o frente a cualquier duda que haya que resolver.



IVAN AYALA REGALADO
Psicólogo Organizacional
Experiencia en coordinación, asesoría, diseño e intervención de factores de riesgo psicosocial.
Gestión y dirección de talento humano
CC 16935556
Cel. 3188268944
Correo-e: ivanayalareg@gmail.com

XVII. Anexo J: Proyecto de investigación llevado actualmente por el investigador en el tema de la tesis doctoral

Actualmente, el investigador se encuentra adelantando un proyecto de investigación de 2 años para la recolección de un conjunto de datos de riesgo psicosocial en la industria de pulpa y papel como parte de un proyecto de investigación del grupo Nuevas Tecnologías Trabajo y Gestión de la Universidad de San Buenaventura, Cali.

	SISTEMA DE GESTIÓN FORMATO ACTA DE INICIO PROYECTO DE INVESTIGACIÓN	
Versión: 1.0	Fecha de Aprobación: 16/JUL/2018	Código: PM.IN.F04

DIRECCIÓN DE INVESTIGACIONES
ACTA DE INICIO PROYECTO DE INVESTIGACIÓN

INFORMACIÓN GENERAL	
FECHA	Septiembre 02 2019
GRUPO DE INVESTIGACIÓN	GRUPO NUEVAS TECNOLOGÍAS, TRABAJO Y GESTIÓN
NOMBRE DEL PROYECTO	Aplicaciones de la Inteligencia Artificial en la Seguridad y Salud en el Trabajo en el reconocimiento de Riesgos Laborales*.
OBJETIVO GENERAL	
Diseñar y desarrollar un sistema de clasificación basado en técnicas inteligentes para identificar los riesgos laborales en empresas Colombianas.	
OBJETIVOS ESPECÍFICOS	
<ol style="list-style-type: none"> 1. Analizar los sistemas de clasificación basados en técnicas inteligentes y no inteligentes que sirvan para diseñar modelos para la identificación, evaluación y control de riesgos laborales en empresas colombianas. 2. Diseñar un método de identificación y control para la prevención de accidentes laborales en población trabajadora colombiana. 3. Validar el sistema de clasificación propuesto con el fin de establecer su efectividad en comparación con los métodos tradicionalmente empleados para el reconocimiento de incidentes y accidentes laborales 	

1. NOMBRE DE LOS INTEGRANTES Y SU VINCULACIÓN

Este documento es propiedad de la Universidad de San Buenaventura Cali. Prohibida su reproducción por cualquier medio, sin previa autorización de la Rectoría de la Universidad.
 Página 1 de 6

 UNIVERSIDAD DE SAN BUENAVENTURA SECCIONAL CALI	SISTEMA DE GESTIÓN FORMATO ACTA DE INICIO PROYECTO DE INVESTIGACIÓN	 SISTEMA DE GESTIÓN DE CALIDAD
Versión: 1.0	Fecha de Aprobación: 16/JUL/2018	Código: PM.IN.F04

Rol dentro del proyecto ¹	Nombre	Tipo de Vinculación Con la USB	Dedicación al proyecto (horas/semana)
Investigador principal	Rodolfo Mosquera Navarro	Profesor Tiempo Completo Asociado	12

2. REDES Y/O GRUPOS VINCULADOS

Nombre de la red	Institución	Nombre investigadores

3. DURACIÓN DEL PROYECTO

TIEMPO (meses)	FECHA DE INICIO	FECHA DE FINALIZACIÓN
	día / mes / año	día / mes / año
15	Septiembre 2019	Diciembre 2020

[Consejo Controlador](#)
 Verificar versión en la Matriz de Control de Información Documentada
 Este documento es propiedad de la Universidad de San Buenaventura Cali. Prohibida su reproducción por cualquier medio, sin previa autorización de la Rectoría de la Universidad
 Página 2 de 6



UNIVERSIDAD NACIONAL DE COLOMBIA

SEDE MANIZALES

2020