

Factores de riesgo asociados al síndrome de *burnout* en docentes de la provincia del Carchi

Risk Factors Associated with Burnout Syndrome in Teachers from the Province of Carchi

<http://doi.org/10.53358/ideas.v7i1.1015>

Juan C. Velasco-Benavides, Erick Herrera

Universidad Politécnica Estatal del Carchi, Posgrado
 Av. Universitaria y Antisana, Tulcán - Carchi

carlos.velasco@upec.edu.ec; erick.herrera@upec.edu.ec

Fecha de envío, febrero 23/2024 - Fecha de aceptación, abril 15/2024 - Fecha de publicación, enero 30/2025

Resumen: El síndrome de burnout es ampliamente conocido como un motivo de problema laboral que aqueja al personal docente a nivel mundial. Este problema relacionado con factores de riesgo que generan estrés surge debido al entorno laboral y diversas situaciones ligadas al trato con personas que demandan un alto nivel de dedicación e involucramiento. El presente estudio se centró en identificar factores de riesgo asociados al síndrome de burnout en docentes de la provincia del Carchi. El objetivo principal fue validar la adaptación del Inventario de Burnout de Maslach (MBI) en la dimensión de cansancio emocional al contexto educativo específico, y proponer un modelo de análisis de resultados basado en inteligencia artificial. Al validar el instrumento se identificó que este no se adaptaba completamente al contexto local por lo que el instrumento fue modificado, teniendo que removerse 5 preguntas que fueron identificadas como no relevantes mediante el Análisis Factorial Confirmatorio, y se incorporó información demográfica de los participantes considerada de interés para las instituciones participantes. De esta manera, para la extracción de resultados a partir del test, se diseñó un nuevo modelo de procesamiento de resultados basado en técnicas multivariantes, contemplando 37 variables de entrada. Para esto se consideraron múltiples alternativas, incluidos clasificadores binarios, redes neuronales poco profundas y un modelo de 5 capas ocultas desarrollado con TensorFlow y Keras con la configuración de 48, 69, 52, 101 y 97 neuronas, que ofreció las predicciones más precisas de entre los modelos evaluados sobre el agotamiento emocional, una dimensión clave del síndrome de burnout. Este modelo alcanzó una precisión del 86% y un MSE de 0.1193604, determinando que los factores de riesgo fueron la disminución y pérdida de recursos emocionales, demostrando su fiabilidad para la detección automática del síndrome sin necesidad de diagnóstico profesional. La investigación valida la adaptación del instrumento al contexto objetivo y resalta la importancia de establecer nuevos modelos de análisis cuando se emplean instrumentos modificados a partir del MBI, asegurando su relevancia y aplicabilidad en contextos específicos.

Palabras Clave: Síndrome de burnout, Factores de Riesgo, Red Neuronal Artificial, Análisis Factorial Confirmatorio, Distancias de Mahalanobis, Accuracy.

Abstract: Burnout syndrome is widely recognized as a significant work-related issue affecting educators worldwide. This problem, associated with risk factors that generate stress, arises from the work environment and various situations involving interactions with individuals requiring a high level of dedication and involvement. This study focused on identifying risk factors associated with burnout syndrome in teachers from the Carchi province. The primary objective was to validate the adaptation of the Maslach Burnout Inventory (MBI) emotional exhaustion dimension to the specific educational context and propose an AI-based results analysis model. The validation process revealed that the instrument did not fully adapt to the local context, leading to its modification by removing five questions identified as irrelevant through Confirmatory Factor Analysis. Demographic information of interest to participating institutions was also incorporated. To extract test results, a new multivariate analysis model was designed, considering 37 input variables. Various alternatives were explored, including binary classifiers, shallow neural networks, and a five-hidden-layer model developed using TensorFlow and Keras, configured with 48, 69, 52, 101, and 97 neurons, which yielded the most accurate predictions for emotional exhaustion—a key dimension of burnout syndrome. This model achieved 86% accuracy and an MSE of 0.1193604, determining that risk factors were the depletion and loss of emotional resources, demonstrating its reliability for automatic burnout detection without requiring professional diagnosis. The study validates the instrument's adaptation to the target context and highlights the importance of establishing new analytical models when using modified MBI instruments, ensuring their relevance and applicability in specific contexts.

Keywords: Burnout Syndrome, Risk Factors, Artificial Neural Network, Confirmatory Factor Analysis, Mahalanobis Distances, Accuracy.

Autor de correspondencia:

Juan Velasco, carlos.velasco@upec.edu.ec

Introducción

El síndrome de burnout asociado a factores de riesgo que provocan estrés en docentes se da por el ambiente laboral y las diferentes situaciones asociadas al contacto con personas que provocan una alta entrega e implicación. La Red Mundial de Salud Ocupacional de la OMS, considera al síndrome de burnout como un desgaste laboral progresivo que tiene su inicio con altos niveles y prolongados de estrés laboral, que se confirman por la irritabilidad, cansancio y tensión. Este síndrome ya es un problema de salud pública debido a los efectos negativos sobre el empleado y el empleador [1].

La Organización Mundial de la Salud (OMS) desde el año 2022 lo considera a nivel mundial como una enfermedad de trabajo y se le conoce como el burnout o síndrome de desgaste ocupacional y que se ha incluido en la nueva Clasificación Internacional de Enfermedades (CIE-11). Por el año 2019 este síndrome era considerado en el listado de Problemas Asociados con el Empleo o el Desempleo, en la actualidad se le reconoce oficialmente como una enfermedad de trabajo. Según la OMS se define como la consecuencia del estrés crónico que se desarrolla donde se trabaja y que no se ha tratado con éxito, especificando que se desarrolla por el contexto laboral y no puede usarse en otras actividades de la vida del ser humano [2].

En la provincia del Carchi donde se desarrolla esta investigación no existe información significativa sobre el síndrome de burnout en docentes, pero si hay información sobre medición del estrés laboral en los docentes, que es uno de los factores de riesgo que sí influyen en que se llegue a desencadenar el síndrome [3]. Por lo cual, se necesita más información que ayude a complementar las investigaciones que ya existen y con ello plantear medidas para enfrentar de forma asertiva los diferentes factores de riesgo que inducen el síndrome de burnout en docentes.

Por la redacción expuesta, el presente trabajo ayudará a los estudios ya existentes que se realizan a nivel nacional, y en particular en la Provincia del Carchi, sobre cuáles son los factores asociados al síndrome de burnout en docentes, como un aporte a las estrategias de cómo enfrentar a este síndrome, mejorar la calidad de vida docente y el manejo correcto de situaciones que influyan en los factores que inciden en el síndrome. Para la predicción del síndrome de burnout según los factores de riesgo asociados en docentes de la provincia del Carchi, se empleó un conjunto de técnicas multivariantes compatibles con la naturaleza de los datos que conforman el instrumento que se usó en este estudio. Como técnicas multivariantes se usaron: clasificadores binarios de una sola capa, redes neuronales poco profundas, y cinco modelos de Deep Learning. De todas las técnicas que se utilizó, la técnica que mejores resultados presentó sobre la predicción del síndrome de burnout en la dimensión de cansancio emocional, fue el Deep Learning con un modelo de cinco capas ocultas a través de TensorFlow y Keras. El clasificador de Deep Learning logró una precisión de 86% y un MSE de 0.1193604, demostrando ser un método confiable para la detección automática del síndrome de burnout en la dimensión de cansancio emocional sin la necesidad de la intervención de un profesional que diagnostique los riesgos asociados al síndrome. El método Deep Learning se adaptó a la población de estudio que fueron los docentes de unidades educativas de la provincia del Carchi, puesto que, el instrumento se modificó al contexto local y se incorporó información demográfica de los participantes considerada de interés para las instituciones participantes.

Trabajos relacionados

El estudio de [4] se realizó en una universidad privada de Madrid, mediante metodología cualitativa de tipo fenomenológica que permitió conocer los factores más importantes del síndrome de burnout en docentes universitarios, desde su propio punto de vista. El estudio se realizó a 25 docentes de la universidad mediante una entrevista semiestructurada enfocada en los factores que producen malestar en el campo laboral. Los docentes tenían edades entre los 35 a 63 años con una media de 52,2 años y desviación estándar de 7,6. Entre los datos sociodemográficos que se consideraron son el estado civil, tipo de contrato y años de experiencia, además del tipo de contrato en la institución. Los resultados muestran que existió una sobrecarga laboral, actividades a desarrollar fuera del lugar de trabajo y la necesidad de tiempo personal; también la percepción de que los alumnos no tienen motivación son algunos de los factores que producen burnout.

Una investigación realizada por [5] en instituciones de educación superior en la ciudad de Guayaquil bajo un análisis cuantitativo a una muestra de 159 docentes y con la técnica de encuesta se encontró como resultados que el síndrome de burnout se diferencia de varias maneras si se considera los años de experiencia en la docencia y según las condiciones del trabajo. Se encontró que el estrés laboral está presente en los docentes con un 88,2% y el síndrome de burnout en un 67,5%, se consideraron niveles altos y reveló la presencia de un problema que se consideró alarmante ya que tiene impacto en la salud, el bienestar y el desempeño, y en efecto su relación con el daño social.

En el artículo de [6] se estudió el burnout en docentes españoles y los factores asociados al síndrome, encontrándose con 52 factores divididos en 7 grupos, la muestra fue de 35 artículos con una muestra de docentes. El instrumento fue el MBI "Maslach Burnout Inventory" de Maslach y Jackson (1989). Los resultados mostraron que las variables más repetitivas son las sociodemográficas con un 94,4%, entre las cuales están el sexo con un 91,34%, la edad con 77,14% y la experiencia docente con 34,29% siendo los factores de mayor interés. Existe una gran variación respecto a los factores revisados en el estudio relacionados con el burnout docente en España, por ello en este estudio se concluyó que es necesario investigar con mayor profundidad para lograr comparaciones más exactas.

Según [7] las variables a analizar fueron organizacionales, personales y demográficas como predictores del síndrome de burnout, se realizó un muestreo estratificado en el municipio de Benidorm, los estratos fueron zonas del municipio y por etapas educativas, en la muestra se consideró 171 docentes de preescolar, primaria y bachillerato. El 72,5% fueron mujeres y el 27,5% hombres. Los resultados mostraron que de acuerdo con el rango de valores entre 1 y 5, los valores obtenidos en agotamiento emocional fueron de 2.21, de 1.67 en despersonalización y 2.07 en la falta de realización personal. Mediante la regresión múltiple las variables organizacionales se diferenciaron de las dimensiones del burnout, respecto a la afectividad negativa implica como factor de vulnerabilidad. Con el agotamiento emocional se tuvieron que el estrés laboral y la baja satisfacción laboral son grandes predictores.

Por su parte, en el estudio de [8] utilizó el método de redes neuronales artificiales con una sola capa oculta (RNA) para investigar la asociación entre varias dimensiones demográficas y de liderazgo de entrenadores con la satisfacción laboral de los docentes en las escuelas coreanas y con un número de neuronas en la capa oculta que varió de 1 a 50. El método que se usó fue un cuestionario estructurado autoadministrado a los maestros de 12 escuelas primarias ubicadas en Ulsan, Corea, obteniendo una muestra de 194 datos que dividieron en 124 datos para el proceso de entrenamiento y desarrollo del modelo y los 70 restantes

para la prueba y validación de la precisión predictiva. El número de capas y neuronas ocultas se seleccionó mediante experimentación. Mediante el RNA se encontró el modelo con mejor capacidad de relación y con mayor precisión predictiva que el análisis de regresión múltiple. Mediante la realización de un software independiente y de fácil manipulación que se usó en términos predictivos y de estimación en búsqueda de la importancia relativa de las variables independientes en la satisfacción laboral. Los resultados mostraron gran complejidad entre las representaciones no lineales que comprenden la relación entre las dimensiones demográficas y de entrenamiento con la satisfacción laboral. Para encontrar la arquitectura óptima del modelo RNA se realizó varios experimentos, un total de 190 arquitecturas RNA. Finalmente, el modelo pudo predecir aproximadamente el 79 % dentro del 20 % de la satisfacción laboral real.

En el estudio de [9] se tuvo como objetivo predecir el burnout profesional y las variables de competencia a 51 profesores de ciencias que trabajan en la escuela secundaria en Zonguldak mediante el uso de una red neuronal artificial RNA. Mediante la red neuronal se buscó explicar los problemas de burnout al considerar el factor humano. La técnica fue la encuesta sobre burnout, autoeficacia y competencias a profesores de ciencias. Con la base de datos se estableció la red neuronal artificial. Los resultados mostraron que la autoeficacia y la competencia de los profesores de ciencias pueden pronosticar burnout profesional en varios grados. Según las dimensiones del burnout que son agotamiento emocional, despersonalización y realización personal se predijo mediante la RNA los siguientes resultados: el rendimiento de la RNA fue del 40 % para la primera dimensión que es el "agotamiento emocional", con un 50 % para "éxito personal", con aproximados del 20 % para "despersonalización" y del 80% para "competencia".

Materiales y métodos

La investigación se realizó en unidades educativas de la provincia del Carchi. La muestra provino de una base de datos de un estudio de pregrado de la Universidad Politécnica Estatal del Carchi de la carrera de enfermería sobre los factores de riesgo relacionados con el síndrome burnout en el periodo 2021-2022 en la cual se analizó a 200 docentes mediante una encuesta.

El enfoque del estudio fue principalmente cuantitativo por las técnicas de recolección de datos que se aplicó, las mediciones y el posterior análisis estadístico de los datos recopilados para medir su frecuencia y determinar la existencia del síndrome de burnout en los docentes. Además, el enfoque cualitativo sirvió en esta investigación para el análisis profundo de los resultados obtenidos de cada modelo según las técnicas multivariantes que se use con los datos para su entrenamiento, permitirá la determinación de ventajas y limitaciones, finalmente seleccionar la mejor alternativa para medir de manera automática la prevalencia de burnout.

El instrumento aplicado se construyó por medio de un cuestionario de dos partes. La primera parte consta de 15 preguntas sobre los datos sociodemográficos, factores individuales, factores sociales, variables sociales extra laborales y variables organizacionales. En la segunda parte se analizó la variable *Síndrome de burnout* para la cual se aplicó el Cuestionario Maslach Burnout Inventory (MBI) donde se consideró a las dimensiones: Agotamiento o cansancio emocional, Despersonalización, Realización personal. El cuestionario de Maslach Burnout Inventory (MBI) está constituido por 22 ítems en forma de afirmaciones, sobre los sentimientos y actitudes del profesional en su trabajo y hacia los alumnos adaptado

especialmente a los docentes y su función es medir el desgaste profesional. Este test permite medir la frecuencia y la intensidad con la que se sufre el síndrome de burnout.

Tratamiento de datos

Distancias de Mahalanobis. Cuando se aplica un instrumento de medición siempre hay la posibilidad de la existencia de datos atípicos y perdidos, por lo cual es recomendable, antes de un análisis estadístico realizar un proceso de aplicación de análisis de datos. Para el tratamiento de datos la técnica más usada es las distancias de Mahalanobis, mediante la desviación estándar de la observación respecto a la media en una distribución. Debido al comportamiento de un dato atípico comparado con las demás observaciones, es una herramienta que permite identificar este tipo de datos. Al comparar la distancia Euclidiana con la de Mahalanobis es la más corta entre dos puntos, pero no considera la correlación entre variables altamente correlacionadas que la distancia de Mahalanobis si toma en cuenta [10], [11]. Esta distancia es una métrica de escala invariante que logra la medida de la distancia entre un punto $x \in \mathbb{R}^p$ que resulta de una distribución de probabilidad $f_x(\cdot)$, p-variada y la media $\mu = E(X)$ de la distribución. Considerando que, la distribución $f_x(\cdot)$ presenta momentos finitos de segundo orden se realiza la matriz de covarianza como $\Sigma = E(X - \mu)(X - \mu)^T$. Las distancias de Mahalanobis se representan como [11]:

$$D(X, \mu) = \sqrt{(X - \mu)^T \Sigma^{-1} (X - \mu)}. \quad (1)$$

Análisis factorial confirmatorio (AFC). Luego de un tratamiento de datos es necesario realizar la validación para seleccionar el conjunto de variables que permitan dar la validez y fiabilidad de cada ítem del cuestionario que se aplicó y del cual se obtuvo la base de datos que se trabaja en el presente estudio. El AFC se considera un método que permite analizar un tipo de variables que son ordinales, relaciona el nivel de las respuestas de un vector $p \times 1$ de variables observadas que son aleatorias, permiten dar explicación de una o más variables que no se observan y se llaman factores η [12]. Este método evalúa y estima en cada ítem su saturación y que explica diferentes aspectos de la variable latente no observada. El siguiente modelo explica al vector Y_i de respuestas observadas que se predicen por la variable latente no observada ξ :

$$Y = \Lambda \xi + \epsilon, \quad (2)$$

donde el vector de dimensión $p \times 1$ es Y , las variables latentes no observadas se representan con ξ y Λ es una matriz de dimensión $p \times k$ en donde k es igual al número de variables latentes no observadas. Además, el modelo considera el error ϵ , porque Y se establece por un grupo de variables que exponen de manera imperfecta a ξ . Para dar solución a este modelo se usa la estimación de máxima verosimilitud (ML), que se logra mediante la reducción iterativa de la función de ajuste:

$$F_{ML} = \ln |\Lambda \Lambda' + I - \text{diag}(\Lambda \Lambda')| + \text{tr}(R(\Lambda \Lambda' + I - \text{diag}(\Lambda \Lambda')^{-1})) - \ln(R) - p \quad (3)$$

Es la matriz de varianza-covarianza que se relaciona en el modelo de AFC propuesto y R es la matriz de varianza-covarianza observada. Esto permite que los parámetros del modelo sean estimados mediante la reducción de la distancia entre la varianza-covarianza relacionada en el modelo y la observada [13].

Análisis de componentes principales (PCA). Es un enfoque estadístico que reduce la complejidad de los espacios muestrales multidimensionales preservando su información y permite ver los datos. Entre sus aplicaciones se tiene la reducción de dimensionalidad (variables), perdiendo la mínima cantidad de información (varianza).

El PCA permite la descripción de tablas que están distribuidas por filas y que constan de las unidades estadísticas llamadas "individuos", y en sus columnas se contiene la información de las variables tipo continuo que se aplicaron como medida hacia los individuos [14]. EL PCA permite calcular el número de capas ocultas que son necesarias para la implementación de una red neuronal. Para una muestra de n individuos cada uno con p variables (X_1, X_2, \dots, X_p), con lo cual la muestra tiene p dimensiones. El número de factores subyacentes que se encuentran con el PCA está definido cuando ($z < p$) y que bajo esta condición se mantiene la misma información original de las p variables. Los componentes principales se representan cada uno con z. Antes de calcular cada componente principal se considera las unidades de medida de las variables. El PCA empieza con la estandarización de las variables para que tengan media 0 y desviación estándar 1:

$$x_i = \frac{x_i - \text{media}(x)}{\text{sd}(x)} \quad (4)$$

Los componentes principales representados por (Z_i) son producto de la combinación lineal entre las variables originales. De un grupo de variables dado por (X_1, X_2, \dots, X_p) la primera componente principal se da como la combinación lineal normalizada de la mayor varianza de las variables.

$$Z_1 = \phi_{11} * X_1 + \phi_{21} * X_2 + \dots + \phi_{p1} * X_p \quad (5)$$

Redes neuronales. Las redes neuronales artificiales son técnicas que permiten el aprendizaje automático que simulan el mecanismo de aprendizaje en organismos biológicos. Este mecanismo biológico se simula en redes neuronales artificiales, que contienen unidades de cómputo denominadas neuronas. Una red neuronal artificial realiza los cálculos al considerar las entradas (x_1, x_2, \dots), al propagar los valores que se calculan desde las neuronas de entrada a las neuronas de salida (Y) y se usa los pesos (w) como parámetros intermedios. El aprendizaje se sujeta al cambio de los pesos que conectan las neuronas. De la base los datos proporcionan la información sobre la exactitud de los pesos en la red neuronal, al comparar sobre qué tan bien coincide la salida pronosticada para una entrada en particular la etiqueta de salida anotada en los datos de entrenamiento [15].

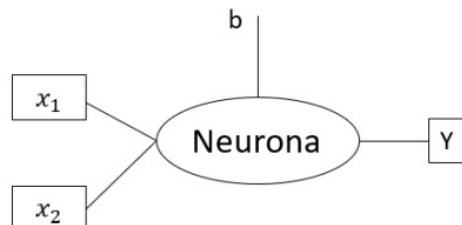


Fig. 1. Esquema principal de una neurona con dos entradas [16].

La salida de la neurona se calcula $Y = x_1 * w_1 + x_2 * w_2 + b$. Como función de activación se seleccionaron la función ReLU que sirve para el diseño de capas ocultas y para la capa de salida la función Softmax que se formula como:

$$f(Z)_j = \frac{e^{z_j}}{\sum_{k=1}^k e^{z_k}} \quad (6)$$

Deep Learning. El Deep learning usa técnicas estadísticas avanzadas como las redes neuronales artificiales, que se basan en capas de nodos interconectados mediante pesos w_{ij} [17]. Mediante el aprendizaje supervisado, minimizaban una función de pérdida como el error cuadrático medio para hacer predicciones sobre datos etiquetados. Esto se logra mediante descenso de gradiente estocástico, actualizando los pesos en la dirección opuesta al gradiente del error. Una fórmula clave es la regla delta para ajustar los pesos [18]:

$$\Delta w_{ij} = \eta (y_j - \hat{y}_j) x_i \quad (7)$$

Donde η es la tasa de aprendizaje. Otras técnicas estadísticas clave son la regularización para evitar overfitting, como el dropout, donde se ignoran aleatoriamente algunas neuronas durante el entrenamiento. Para generar predicciones se utiliza la propagación hacia adelante mediante la fórmula:

$$\hat{y} = f(X; w) \quad (8)$$

Donde \hat{y} es la predicción, X los inputs y f la función de activación no lineal como ReLU aplicada en cada capa.

Curva ROC. La curva ROC (Receiver Operating Characteristic) se utiliza como una herramienta para evaluar la eficiencia de los modelos de clasificación binaria. A través de esta curva, se midió el rendimiento de los modelos y se determinó su capacidad para distinguir entre las clases positivas y negativas. La eficiencia del modelo mediante la curva ROC se evalúa utilizando el área bajo la curva (AUC, por sus siglas en inglés).

El AUC representa la probabilidad de que el modelo clasifique correctamente una instancia positiva con una puntuación más alta que una instancia negativa. Cuanto mayor sea el AUC, más eficiente será el modelo en la tarea de clasificación binaria. Un AUC superior a 0.8 se considera un buen rendimiento, mientras que un AUC superior a 0.9 se considera excelente [19].

Accuracy. En el contexto del aprendizaje profundo (Deep learning), la precisión (accuracy) es una métrica estadística que mide la capacidad de un modelo para realizar predicciones correctas sobre un conjunto de datos. Representa la proporción de instancias que el modelo ha clasificado correctamente en comparación con el total de instancias. Se calcula mediante la fórmula:

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

La precisión se calcula dividiendo el número de predicciones correctas entre el número total de predicciones realizadas por el modelo. Es una métrica ampliamente utilizada debido a su simplicidad e interpretabilidad directa [20].

Error medio cuadrático. El error cuadrático medio (MSE, por sus siglas en inglés) es una métrica de evaluación ampliamente utilizada en el aprendizaje profundo (Deep learning) para medir el rendimiento de los modelos en tareas de regresión. Esta métrica cuantifica la diferencia promedio entre las predicciones realizadas por el modelo y los valores reales observados. El MSE se calcula mediante la fórmula:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (10)$$

Un valor más bajo del MSE indica un mejor ajuste del modelo a los datos, ya que las predicciones se encuentran más cercanas a los valores reales. Sin embargo, es importante tener en cuenta que el MSE puede ser sensible a valores atípicos (outliers) en los datos, ya que el cuadrado de las diferencias grandes aumentará significativamente el valor del MSE [21].

Overfitting. El sobreajuste (Overfitting) es un problema común en el aprendizaje profundo (Deep learning) que ocurre cuando un modelo se ajusta demasiado a los datos de entrenamiento, aprendiendo incluso los patrones irrelevantes presentes en esos datos. Esto conduce a un rendimiento deficiente cuando el modelo se utiliza con nuevos datos que no ha visto antes. Un modelo sobre ajustado tiene una alta precisión en los datos de entrenamiento, pero un rendimiento pobre en los datos de prueba o validación. Esto se debe a que el modelo ha "memorizado" las características específicas de los datos de entrenamiento en lugar de aprender los patrones generales subyacentes que permiten una buena generalización a nuevos datos. Para evitar el sobreajuste, existen varias técnicas que se pueden aplicar durante el entrenamiento, como la regularización (como la regularización L1 o L2), el dropout, el aumento de datos (data augmentation) y el uso de conjuntos de validación para detener el entrenamiento cuando el rendimiento en estos conjuntos comienza a disminuir [22].

Resultados

La base de datos consta de 200 observaciones que provienen del instrumento aplicado que consta de 37 preguntas categóricas distribuidas en dos secciones. La primera parte de 15 preguntas sobre los datos sociodemográficos, factores individuales, factores sociales, variables sociales extra laborales y variables organizacionales y la segunda parte de 22 preguntas sobre la variable *síndrome de burnout*, medida con el cuestionario Maslach Burnout Inventory (MBI). El instrumento aplicado es multivariado categórico y ordinal, por lo cual las técnicas estadísticas convencionales no son suficientes para estudiar el problema planteado, con lo cual fue necesario el uso de redes neuronales artificiales y el Deep Learning que son herramientas ampliamente utilizadas en los últimos años para resolver problemas complejos [23].

La base de datos se trabajó con el lenguaje de programación estadística R, además con Python con su distribución Anaconda para el manejo de los paquetes TensorFlow y Keras desde el programa RStudio, con el uso de la librería reticulate. El análisis de la base de datos se inició con la imputación mediante el uso de la técnica KNN (K-Nearest Neighbors) y la librería VIM, los resultados muestran que no existen datos faltantes.

A continuación, se hizo el tratamiento a los datos para detectar datos atípicos con el uso de las Distancias de Mahalanobis. Se inició con el puntaje de corte basado en la distribución χ^2 y con un 99.9% de las distancias con lo cual se excluyó al 0.1% de los cuantiles de las observaciones que se consideraron como atípicas [24]. El resultado demuestra que no se excluyó ninguna observación atípica y por lo tanto la muestra sigue compuesta por 200 observaciones.

Antes de trabajar el instrumento fue necesaria la validación previa a la obtención de inferencias y resultados mediante la base de datos. Se utilizó el Análisis Factorial Confirmatorio (AFC) [25], esta técnica permite comprobar que tan válido y fiable es un instrumento. Como técnica paramétrica se realizó 5 supuestos a verificar como son: aditividad, normalidad, linealidad, homogeneidad y homocedasticidad. Para la aditividad se realizó la matriz de correlación multivariada con cada par posible de preguntas, el objetivo es demostrar que ningún par de preguntas presenta correlación perfecta. La matriz de correlación multivariada para las preguntas del test de burnout en el supuesto de aditividad se muestra en la figura 2 en el siguiente enlace: <https://goo.su/vXhB>.

La figura 2 muestra que ningún par de preguntas presentan correlación alta (0.95 a 1), con lo cual ninguna pregunta tiene correlación perfecta, por lo que se aceptó el supuesto de aditividad.

Para los restantes supuestos paramétricos, se empleó la falsa regresión, mediante un set de cuantiles aleatorios que se basan en la distribución X^2 , los cuales fueron escalados, estudentizados y se ajustan por regresión para luego comparar con los cuantiles de la muestra. La normalidad multivariada se analizó con el histograma de los cuantiles obtenidos con la falsa regresión, el supuesto de linealidad se realizó con el Q-Q Plot y la homogeneidad y homocedasticidad, se analizaron con el Scatterplot [26]. El análisis y resultados de la falsa regresión se muestran en la figura 3 y se encuentran en: <https://goo.su/vXhB>.

En la figura 3a, los cuantiles ajustados tienen una distribución similar a la normal, con lo cual se aceptó la normalidad. En la figura 3b, se observó que los cuantiles teóricos respecto a los muestrales tienen una distribución cercana a la tendencia lineal, con lo cual se aceptó el supuesto de linealidad. Finalmente, en la figura 3c, se mostró que los cuantiles estandarizados respecto a los ajustados tienen una distribución homogénea en los cuatro cuadrantes, no se apreció patrones o agrupamientos inadecuados, lo que se observó es que los cuantiles se distribuyeron en forma de salpicado aleatorio, con lo cual se aceptó la homogeneidad y homocedasticidad.

Los 5 supuestos han sido verificados y por lo tanto se procedió a realizar el AFC que permite dar la validez y fiabilidad del instrumento aplicado, cada pregunta responde a un factor y aporta a la estructura factorial. Por lo cual los factores que se utilizaron para el análisis fueron:

Tabla 1. Estadísticos descriptivos para las variables ordinales en la base de datos original.

Factor	Preguntas
Cansancio emocional	p1, p2, p3, p6, p8, p13, p14, p16, p20
Despersonalización	p5, p10, p11, p15, p22
Realización personal	p4, p7, p9, p12, p17, p18, p19, p21

El análisis factorial confirmatorio se realizó en el programa estadístico RStudio, con el uso de las librerías lavaan y SemPlot. Por lo tanto, la configuración de la estructura factorial, las saturaciones y sus correlaciones inter e intrafactoriales quedan de la siguiente manera (Fig. 4.) y se encuentran en: <https://goo.su/vXhB>.

En la figura 4, el Path-diagram del AFC muestra las correlaciones inter e intrafactoriales y las saturaciones de cada ítem del test de burnout con la estructura factorial. El AFC menciona que la saturación no debe tener valores debajo de 0.3 y también ningún par de factores deben presentar correlación perfecta [10]. Como se observó en la figura 4, existen algunos ítems con valores menores a los recomendados por lo cual se retiraron las preguntas p10,p15,p18,p19 y p21. Luego se volvió a correr el código y no se evidenció indicios de invalidez en el modelo. Los resultados obtenidos mediante el AFC se presentan en la tabla 2 y los índices de bondad de ajuste globales en la tabla 3 y que se encuentran en: <https://goo.su/jC6z>.

Tal como se observó en la tabla 2, todos los ítems considerados en el modelo alcanzan niveles significativos, por lo que no es necesario modificar nuevamente el modelo del test de burnout. Al analizar los valores obtenidos se apreció que los ítems contribuyen en distinta medida dentro de su respectivo factor. En la tabla 3, los índices de bondad de ajuste muestran que el Índice de Ajuste Comparativo (CFI), el Índice de Tucker-Lewis (TLI) y el Índice de Ajuste No Normado (NNFI) alcanzaron valores de 0.9, mientras que el Error Cuadrático Medio de Aproximación (RMSEA) y el Residual Estandarizado Cuadrático Medio (SRMR) presentan valores inferiores a 0.05. Estos resultados indican que el instrumento cuenta con validez y confiabilidad adecuadas.

Con la base de datos correctamente tratada y validada es necesario dar el puntaje a cada factor de acuerdo con la teoría que define al test de burnout. Como se observó en la figura 4, existen algunos ítems con valores menores a los recomendados según el AFC ya que se menciona que la saturación no debe tener valores debajo de 0.3 y también ningún par de factores deben presentar correlación perfecta por lo cual se retiraron las preguntas p10, p15, p18, p19 y p21.

Por lo tanto, los puntajes que definen a cada factor y su categorización acorde con la escala se definieron en la tabla 4.

Tabla 2. Puntajes y escala para cada factor.

Factor	Puntaje	Criterio
Cansancio emocional	$CnE = p1 + p2 + p3 + p6 + p8 + p13 + p14 + p16 + p20$	Bajo: $0 \leq CnE \leq 18$ Medio: $19 \leq CnE \leq 26$ Alto: $CnE \geq 27$
Despersonalización	$Dsp = p5 + p11 + p22$	Bajo: $0 \leq Dsp \leq 3$ Medio: $4 \leq Dsp \leq 5$ Alto: $Dsp \geq 6$
Realización personal	$RIP = p4 + p7 + p9 + p12 + p17$	Bajo: $RIP \leq 24$ Medio: $22 \leq RIP \leq 24$ Alto: $RIP \geq 25$

La tabla 4 presentó la clasificación de los distintos factores evaluados mediante el test de burnout. Específicamente, mostró los ítems del test que alcanzaron una saturación adecuada para cada factor. Además, indica los puntos de corte utilizados para categorizar los niveles de burnout como bajo, medio o alto en cada factor.

A continuación, en primer lugar, las 15 preguntas iniciales del cuestionario, de tipo categórico sobre datos sociodemográficos y otros factores, fueron transformadas en variables dummy mediante la codificación de niveles correspondiente con las librerías `recipes` y `tidyverse`. Con lo cual, la base de datos codificada quedó constituida por 98 variables. De ellas, 89 se consideraron como variables predictoras (features) para la capa de entrada neuronal. Las variables `CE_bajo`, `CE_medio`, `CE_alto`, `DP_bajo`, `DP_medio`, `DP_alto`, `RP_bajo`, `RP_medio` y `RP_alto` (Cansancio Emocional, Despersonalización y Realización Personal) se definieron como variables respuesta (labels). Posteriormente, se utilizaron las librerías `GGally` y `skimr` para la visualización y verificación de los datos antes del entrenamiento de modelos. Los resultados de los estadísticos descriptivos de las variables codificadas de este procesamiento y transformación de variables se muestran en la tabla 5 en el siguiente link: <https://goo.su/jC6z>.

La tabla 5 muestra que al realizar un adecuado tratamiento de los datos se obtuvo una base de datos sin valores faltantes ni atípicos, donde todas las variables presentaron una varianza distinta de cero. Como se observó, cada variable tiene una dispersión y distribución diferente. Por esto, se diseñó una capa de normalización a la entrada que se ajustó al comportamiento de los datos, permitiendo a la red neuronal trabajar con escalas similares y evitando así que la escala de los gradientes utilizados durante el entrenamiento se afecte. La capa de normalización de entrada se desarrolló utilizando las funciones `layer_normalization` y `adapt` de Keras [27]. Las variables de respuesta también fueron codificadas en formato dummy, por lo que se diseñó una capa de salida con tres neuronas, ya que para el síndrome de burnout se midieron respuestas en los niveles bajo, medio y alto. La codificación se realizó con la función `to_categorical` de Keras.

Como se tiene tres variables de respuesta en el presente estudio se trabajó solo uno de los casos como lo es la variable Cansancio Emocional.

Análisis del Factor Cansancio Emocional

El test de burnout fue creado con el fin de permitir evaluar el nivel de cansancio emocional, despersonalización y realización personal de los individuos que participaron en el experimento. Sin embargo, de estos tres factores, el presente estudio se enfocó en el factor cansancio emocional que es de especial interés para las instituciones educativas donde laboran los docentes de la población estudiada. De esta manera, en el presente estudio se continuó con la obtención de un modelo de inteligencia artificial que permitió evaluar el cansancio emocional de los participantes mediante el modelo validado y aplicable en la población de docentes de la provincia del Carchi. Cabe destacar que, el modelo original está compuesto de 22 preguntas, sin embargo, mediante la muestra evaluada en este estudio se determinó que 5 preguntas no saturaron la estructura factorial, por lo que fueron removidas, de manera que el modelo quedó conformado por 17 preguntas, las cuales requieren de un nuevo mecanismo que permitió evaluar el cansancio emocional en el docente. Para esto se desarrolló un nuevo modelo de inteligencia artificial capaz de clasificar a los docentes en tres niveles de cansancio emocional considerando la información completa del test de burnout. Se seleccionó las redes neuronales como alternativa principal

debido a la naturaleza no lineal de los datos ya que este tipo de modelos permiten tomar la totalidad de información del instrumento como variables de entrada además de la información categórica descrita en la tabla 5.

Como parte inicial se desarrolla una red neuronal artificial para la clasificación, pero sin capas ocultas, la cual consta de una capa de normalización y 3 neuronas en la capa de salida. El modelo se entrenó con 243 epochs de aprendizaje, mediante la optimización por Stochastic Gradient Descent (SGD), con Momentum de valor 0.8, y un learning rate decay inicial de 0.1, el cual decrece en una razón de $0.1/243$ para cada etapa de aprendizaje nueva. El valor de 243 epochs de aprendizaje viene por la rule of thumb de [28], que permite el triple del número de variables para las etapas de aprendizaje. Los resultados de esta etapa de aprendizaje se muestran en la figura 5 y su arquitectura para el clasificador en la figura 6 y se pueden visualizar en: <https://goo.su/FhIDwBC>.

El modelo de clasificación diseñado con tres neuronas en la capa de salida y sin capas ocultas, fue evaluado en el conjunto de prueba que corresponde al 10% de las observaciones totales (20 observaciones que el clasificador no había visto previamente). Al evaluar el desempeño del clasificador con estas nuevas observaciones, se obtuvo un Loss de 4.3403816, Accuracy de 0.5000000 y un MSE de 0.2926763

Determinación de la topología de la red neuronal.

Como se observó en los resultados del clasificador de la Figura 6, las métricas de desempeño aún estuvieron lejos de un rendimiento óptimo. Por esto, se propuso un conjunto de modelos de redes neuronales poco profundas y redes neuronales profundas para mejorar el desempeño del clasificador. Para determinar la topología óptima de la red neuronal se utilizó una técnica compuesta por: un análisis de componentes principales (PCA) que permitió encontrar el número óptimo de capas ocultas de la red neuronal [29], y la exploración de todas las posibles configuraciones del número de neuronas en las capas ocultas siguiendo las recomendaciones de [28].

El PCA como técnica de reducción de dimensiones, permitió calcular el número de variables mediante las cuales se pudo determinar progresivamente la varianza de un conjunto de variables. Los resultados del análisis de componentes principales ejecutado sobre la base de datos se ejecutaron mediante la función princomp de R y se muestran en la tabla 6 y figura 7 en el siguiente enlace: <https://goo.su/FhIDwBC>.

Como se observó en la figura 7, se necesitó más de 4 componentes principales en el modelo para poder explicar más del 70% de la varianza presente en los datos. Por lo cual y de acuerdo con [30], se probaron modelos de hasta 5 capas ocultas para configurar la topología óptima de la red neuronal.

Modelo de red neuronal con una capa oculta.

Para determinar el número óptimo de neuronas con una capa oculta del modelo de red neuronal poco profunda, se diseñó un bucle con varias redes neuronales que almacenaban los parámetros y métricas de desempeño. La función de activación utilizada en esta capa fue la rectificadora lineal (ReLU), con regularización L2 y penalización $L=0.001$ para evitar el sobreajuste al aumentar el número de neuronas. Nuevamente se utilizó descenso de gradiente estocástico (SGD) con una tasa de aprendizaje de 0.1 y momentum de 0.8. Considerando las 81 variables de entrada, el intervalo de posibles configuraciones de

neuronas evaluado fue de 40 a 162, es decir, entre la mitad y el doble de las variables de entrada. Los resultados de las métricas de desempeño para la mejor configuración de red neuronal con una sola capa oculta se muestran en la tabla 7 en el link: <https://goo.su/FhIDwBC>.

En la tabla 7 se apreció varias configuraciones de número de neuronas con una sola capa oculta para el modelo de red neuronal, existen algunas configuraciones que presentan un mayor desempeño. Se seleccionó la configuración de 48 neuronas la cual alcanzó las mejores métricas del bucle iterativo que se realizó para encontrar la mejor configuración. El proceso de entrenamiento y arquitectura de la red neuronal con una sola capa oculta se observan en el enlace <https://goo.su/FhIDwBC> y en las figuras 8 y 9.

Modelo de Deep Learning con dos capas ocultas.

Para continuar explorando el modelo, se añadió una segunda capa oculta con el fin de evaluar si esto genera mejoras en el rendimiento en comparación con la configuración previa. En la primera capa oculta se determinó que 48 neuronas ofrecían el mejor desempeño. Por lo tanto, en esta nueva capa oculta se probó combinaciones que iban desde 24 hasta 96 neuronas, es decir, desde la mitad hasta el doble de neuronas que en la primera capa. Se mantuvo el uso de la función de activación ReLU, regularización L2 y penalización $L=0.001$, así como la optimización mediante SGD con tasa de aprendizaje de 0.1 y momentum de 0.8. Los resultados de las métricas de rendimiento para las distintas configuraciones de la segunda capa oculta se presentan en la tabla 8 (<https://goo.su/FhIDwBC>).

En la tabla 8 se apreció varias configuraciones de número de neuronas para el modelo con dos capas ocultas, existen algunas configuraciones que presentan un mayor desempeño. Se seleccionó la configuración de 69 neuronas la cual alcanzó las mejores métricas del bucle iterativo que se realizó para encontrar la mejor configuración. El proceso de entrenamiento y arquitectura de la red neuronal con dos capas ocultas se observan en las figuras 10 y 11 y se encuentran en: <https://goo.su/FhIDwBC>.

Modelo de Deep Learning con tres capas ocultas.

Se volvió a ejecutar el proceso anterior, se añadió una tercera capa oculta con el fin de evaluar si esto genera mejoras en el rendimiento en comparación con la configuración previa. En la segunda capa se determinó 69 neuronas para el mejor desempeño. En la tercera capa oculta se evaluó cada modelo posible desde 34 hasta 138 neuronas, es decir, desde la mitad hasta el doble de neuronas de la capa anterior. Se mantuvo el uso de la función de activación ReLU, regularización L2 y penalización $L=0.001$, así como la optimización mediante SGD con tasa de aprendizaje de 0.1 y momentum de 0.8. Los resultados obtenidos para las distintas configuraciones de la tercera capa oculta se presentan en la tabla 9 (<https://goo.su/FhIDwBC>).

En la tabla 9 se observó varias configuraciones de número de neuronas para el modelo con tres capas ocultas, existen algunas configuraciones que presentaron un mayor desempeño. Se seleccionó la configuración de 52 neuronas la cual alcanzó las mejores métricas del bucle iterativo que se realizó para encontrar la mejor configuración. El proceso de entrenamiento y arquitectura de la red neuronal con tres capas ocultas se observan en las figuras 12 y 13 y se pueden observar en el link: <https://goo.su/FhIDwBC>.

Modelo de Deep Learning con cuatro capas ocultas.

A continuación, se repitió el proceso anterior, se añadió una cuarta capa oculta con el fin de evaluar si esto genera mejoras en el rendimiento en comparación con la configuración previa. En la tercera capa se determinó 52 neuronas para el mejor desempeño. En esta cuarta capa oculta se evaluó cada modelo posible desde 26 hasta 104 neuronas. Esta capa se configuró igual que las anteriores con función de activación ReLU, regularización L2, penalización, optimización mediante SGD, tasa de aprendizaje y momentum. Los resultados obtenidos para las distintas configuraciones de la tercera capa oculta se presentan en la tabla 10 y se encuentran en: <https://goo.su/FhIDwBC>.

En la tabla 10 se evidenció varias configuraciones de número de neuronas para el modelo con cuatro capas ocultas, existen algunas configuraciones que presentan un mayor desempeño. Se seleccionó la configuración de 101 neuronas la cual alcanzó las mejores métricas del bucle iterativo que se realizó para encontrar la mejor configuración. El proceso de entrenamiento y arquitectura de la red neuronal con cuatro capas ocultas se observan en las figuras 14 y 15 (<https://goo.su/FhIDwBC>)

Modelo de Deep Learning con cinco capas ocultas.

Finalmente, se determinó la configuración óptima para un modelo de red neuronal con cinco capas ocultas. En esta capa igualmente se determinó el rango de 50 a 202 neuronas de las posibles configuraciones. Igual que las capas anteriores se configuró con función de activación ReLU, regularización L2, penalización, optimización mediante SGD, tasa de aprendizaje y momentum. Los resultados obtenidos para las distintas configuraciones de la tercera capa oculta se presentan en la Tabla 11.

Tabla 11. Métricas de desempeño para diferentes configuraciones de redes neuronales con cinco capas ocultas.

Number of neurons	loss	accuracy	mse
92	2.2991545	0.6500000	0.2124649
93	2.7440844	0.6000000	0.2269714
94	2.1538401	0.6500000	0.2225322
95	2.2701786	0.6000000	0.2240053
96	2.4056532	0.6000000	0.2503015
97	1.3013562	0.8600000	0.1193604
98	2.2806649	0.5000000	0.2424285
99	2.3120291	0.6000000	0.2067147
100	2.3737655	0.6500000	0.2105785
101	2.5503664	0.6500000	0.2222144
102	3.6234131	0.5500000	0.2649572

Como se observó en la tabla 11, la mejor configuración para el modelo de cinco capas ocultas se obtuvo con 97 neuronas. Comparado con los modelos y capas anteriores las métricas con los mejores resultados de loss, accuracy y mse se obtuvieron en esta última capa. El proceso de entrenamiento y arquitectura de la red neuronal con cinco capas ocultas se observan en las figuras 16 y 17.

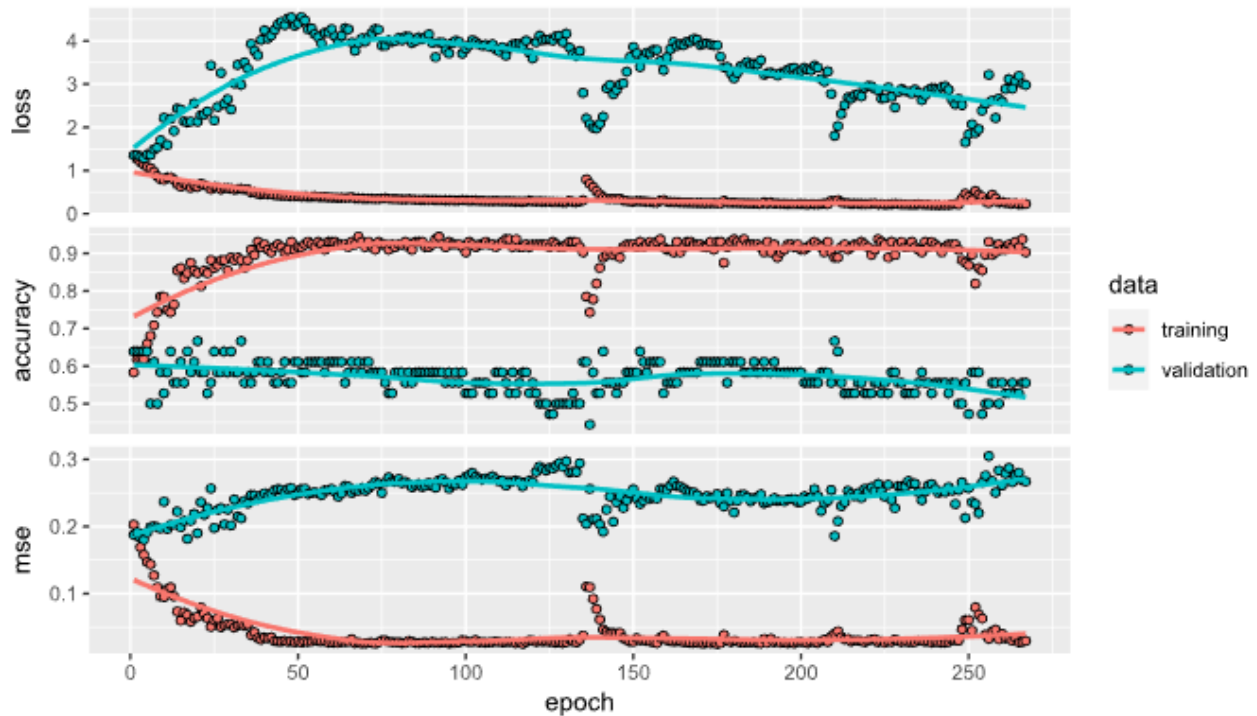


Fig. 16. Proceso de entrenamiento de la red neuronal propuesta con cinco capas ocultas.

Model: "sequential_614"

Layer (type)	Output Shape	Param #	Trainable
normalization_3 (Normalization)	(None, 81)	163	Y
dense_2447 (Dense)	(None, 48)	3936	Y
dense_2446 (Dense)	(None, 69)	3381	Y
dense_2445 (Dense)	(None, 52)	3640	Y
dense_2444 (Dense)	(None, 101)	5353	Y
dense_2443 (Dense)	(None, 97)	9894	Y
dense_2442 (Dense)	(None, 3)	294	Y

=====
 Total params: 26,661
 Trainable params: 26,498
 Non-trainable params: 163
 =====

Fig. 17. Arquitectura de la red neuronal propuesta con cinco capas ocultas.

Finalmente, se verificó cada modelo con su desempeño con los datos de prueba que no se observaron en cada modelo durante el proceso de entrenamiento. Este conjunto estuvo conformado por 20 observaciones y se midió su desempeño con las métricas Loss, Accuracy y MSE. Los resultados se muestran en la siguiente tabla 12.

Tabla 12. Comparación de los modelos implementados, evaluados sobre la base datos Test.

Métricas	Clasificador sin capas ocultas	Modelo secuencial de una capa oculta	Modelo secuencial de dos capas ocultas	Modelo secuencial de tres capas ocultas	Modelo secuencial de cuatro capas ocultas	Modelo secuencial de cinco capas ocultas
Loss	4.3403816	1.0896268	1.9557455	1.4989994	1.7015085	1.3013562
Accuracy	0.5000000	0.7500000	0.7500000	0.7500000	0.7500000	0.8600000
MSE	0.2926763	0.1449158	0.1577028	0.1464316	0.1537883	0.1193604

Como se observó en la Tabla 12, el mejor modelo fue con cinco capas ocultas y 97 neuronas, llegando a un desempeño con las métricas de Loss=1.3013562, Accuracy=0.8600000 y MSE= 0.1193604.

Este modelo fue examinado más a fondo utilizando la matriz de confusión y la curva ROC. Al hacerlo, se logró un rendimiento del modelo de 95.2% y una precisión (Accuracy) de 0.860 en la clasificación del nivel medio de cansancio emocional. Los resultados del rendimiento del clasificador de Deep Learning con 5 capas ocultas se muestran en las figuras siguientes.

Confusion Matrix y Curva ROC

Las imágenes de las matrices de confusión y Curvas ROC para el clasificador de Deep Learning con 5 capas ocultas en los niveles de cansancio emocional alto, medio y bajo se encuentran en: <https://goo.su/fW5Motk>

Tabla 13. Comparación del rendimiento del modelo (AUC), accuracy y precisión para el clasificador de Deep Learning con 5 capas ocultas

Métricas	Cansancio Emocional Alto	Cansancio Emocional Medio	Cansancio Emocional Bajo
AUC	56.1%	95.2%	87.6%
Accuracy	0.45	0.86	0.4
Precision	0.692	1	0.308

Discusión

Los resultados evidenciaron que existen diversas técnicas multivariantes propuestas en varios textos para analizar variables categóricas. Al codificar de forma binaria todas las variables (37 variables categóricas que resultaron en 98 variables dummy), se observó que técnicas convencionales como árboles de decisión, regresiones múltiples y logísticas no son lo suficientemente robustas para obtener buenos resultados con los distintos modelos, dado el conjunto de datos propuesto. Por lo contrario, las redes neuronales son ensamblajes de neuronas artificiales, en donde cada una tiene la posibilidad de aprender los comportamientos no lineales partiendo de los datos. Como se pudo visualizar en la Tabla 12, los modelos de Deep Learning alcanzan niveles de desempeño y precisión excelentes en la tarea de encontrar el cansancio emocional en los docentes de la provincia del Carchi. El modelo óptimo fue de 5 capas ocultas el cual alcanzó una precisión del 86,0% y una sensibilidad del 76,5% para cansancio emocional nivel medio, métricas que demuestran el alto desempeño y potencia que poseen las técnicas de inteligencia artificial para actividades de diagnóstico. Otra de las ventajas que presenta el modelo propuesto es que no invade la integridad de las personas, en consecuencia, la implementación del modelo propuesto puede ser usado como referencia para otras instituciones educativas que requieran evaluar el nivel de cansancio emocional en los docentes.

Al relacionar la información del estudio previo de [5] en donde la investigación se realizó en una provincia diferente a la de nuestro estudio y únicamente con una encuesta que no considera los datos sociodemográficos, factores individuales, factores sociales, variables sociales extra laborales y variables organizacionales como lo hace el presente estudio, además de considerar para la base de datos todos los factores que mide el Cuestionario Maslach Burnout Inventory (MBI), que está constituido por 22 ítems en forma de afirmaciones sobre los sentimientos y actitudes del profesional en su trabajo y hacia los alumnos adaptado especialmente a los docentes y su función es medir el desgaste profesional, en el modelo propuesto de este estudio se puede aplicar basándose en datos categóricos que representen factores de riesgo al relacionarse con factores que previamente fueron identificados. Con lo cual es posible realizar un diagnóstico preciso del riesgo de sufrir burnout en la dimensión de cansancio emocional, esto limitado al grupo de docentes de las instituciones que formen parte de la base de datos, con lo cual es posible tomar medidas correctivas frente al resultado. El presente método puede ser complementado con apoyo psicológico, que permita identificar de manera personal en cada docente cuáles son los factores que les están afectando.

Conclusiones

El presente estudio, enfocado en docentes de la provincia del Carchi, buscó identificar factores de riesgo asociados al burnout y validar el Inventario de Burnout de Maslach (MBI) para un contexto educativo específico, adaptándolo mediante la eliminación de cinco ítems que fueron identificados como no aplicables mediante la técnica de Análisis Factorial Confirmatorio y la inclusión de información demográfica de los participantes. Para el procesamiento de la información extraída mediante el test, se diseñó un nuevo modelo basado en técnicas multivariantes. Este modelo se desarrolló con base a la información obtenida de los datos sociodemográficos, factores individuales, factores sociales, variables sociales extra laborales, variables organizacionales y del Cuestionario Maslach Burnout Inventory (MBI) que están asociados al síndrome de burnout. El instrumento constó de 37 variables que consideran los factores anteriormente mencionados. Todos los datos obtenidos de cada docente fueron estructurados como observaciones las cuales ayudaron a determinar los clasificadores automáticos obtenidos por técnicas multivariantes. Entre los clasificadores que se usaron en el presente estudio estuvieron los clasificadores neuronales sin capas ocultas, redes neuronales poco profundas y varios modelos de Deep Learning. De todas las técnicas implementadas se determinó que el modelo de 5 capas ocultas con la configuración de 48, 69, 52, 101 y 97 neuronas fue el que logró el mejor desempeño para la detección automática del síndrome de burnout en la dimensión de cansancio emocional, llegando a un desempeño con las métricas de Loss=1.3013562, Accuracy=0.8600000 y MSE= 0.1193604, considerando la existencia de una base de datos de prueba que no fue observada por el clasificador durante el proceso de entrenamiento.

Se concluye entonces que existe la posibilidad de la detección automática del síndrome de burnout en la dimensión de cansancio emocional en docentes de unidades educativas de la provincia del Carchi, permitiendo además, considerar factores demográficos y todos los factores asociados al síndrome de burnout de manera precisa y confiable, mediante el uso de técnicas de Deep Learning que siguen mostrando su alto potencial en el diagnóstico de riesgos asociados a síndromes en profesionales de la educación.

Referencias

- [1]. C. Lauracio and T. Lauracio, "Síndrome de Burnout y desempeño laboral en el personal de salud," *Revista Innova Educación*, vol. 2, no. 4, pp. 543–554, Oct. 2020, doi: 10.35622/j.rie.2020.04.003.
- [2]. D. Abril, G. Cabrera, M. Cajamarca, and E. Gutiérrez, "Sistema informático para gestionar información del Síndrome de Burnout en enfermería durante la COVID-19," vol. 15, 2022, [Online]. Available: <http://publicaciones.uci.cu>
- [3]. K. Bolaños, "Medición del estrés laboral en los docentes de la Unidad Educativa José Julián Andrade Carchi Ecuador 2020," 2021.
- [4]. M. P. Egea-Romero and M. I. Galindo-Abradelo, "Factores de riesgo para padecer burnout en docentes universitarios: un estudio cualitativo," *VISUAL REVIEW. International Visual Culture Review / Revista Internacional de Cultura Visual*, vol. 9, no. Monográfico, pp. 1–11, 2022, doi: 10.37467/revvisual.v9.3635.
- [5]. Á. Rivera Guerrero, P. Segarra Jaime, and G. Giler Valverde, "Síndrome de Burnout en docentes de instituciones de educación superior," *Archivos Venezolanos de Farmacología y Terapéutica*, vol. 38, no. 2, pp. 17–23, 2018.
- [6]. C. Pinel-Martínez, M. D. C. Pérez-Fuentes, and J. J. Carrión-Martínez, "Researching burnout in spanish teachers: A review of the associated factors and assessment instruments," *Bordon, Revista de Pedagogia*, vol. 71, no. 1, pp. 115–131, 2019, doi: 10.13042/Bordon.2019.62122.
- [7]. J. Esteras, P. Chorot, and B. Sandín, "Predicción del burnout en los docentes: Papel de los factores organizacionales, personales y sociodemográficos," *Revista de Psicopatología y Psicología Clínica*, vol. 19, no. 2, pp. 79–92, 2014, doi: 10.5944/rppc.vol.19.num.2.2014.13059.
- [8]. B. W. Seok, K. hoan Wee, J. young Park, D. Anil Kumar, and N. S. Reddy, "Modeling the teacher job satisfaction by artificial neural networks," *Soft comput*, vol. 25, no. 17, pp. 11803–11815, 2021, doi: 10.1007/s00500-021-05958-0.
- [9]. I. Özdemir and D. Polat, "International Journal of Educational Studies Available Online at ESci Journals AWARENESS AND RELATIONSHIP MANAGEMENT) AND LEADERSHIP," *International Journal of Educational Studies*, vol. 04, no. 02, pp. 37–47, 2017, [Online]. Available: <https://esciencepress.net/journals/index.php/IJES/article/view/1935/1126>
- [10]. A. E. Jácome Ortega, J. A. Caraguay Procel, E. P. Herrera-Granda, and I. D. Herrera Granda, "Confirmatory Factorial Analysis Applied on Teacher Evaluation Processes in Higher Education Institutions of Ecuador," 2020, pp. 157–170. doi: 10.1007/978-3-030-37221-7_14.
- [11]. H. Ghorbani, "MAHALANOBIS DISTANCE AND ITS APPLICATION FOR DETECTING MULTIVARIATE OUTLIERS," *Facta Universitatis, Series: Mathematics and Informatics*, p. 588, Oct. 2019, doi: 10.22190/FUMI1903583G.

- [12]. E. P. Herrera-Granda, L. L. Lorente-Leyva, J. Yambay, J. Aranguren, M. Ibarra, and J. Peña, "Controller Modeling of a Quadrotor," *Ingénierie des systèmes d'information*, vol. 27, no. 1, pp. 21–28, Feb. 2022, doi: 10.18280/isi.270103.
- [13]. M. Martínez Ávila, "Análisis factorial confirmatorio: un modelo de gestión del conocimiento en la universidad pública," *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, vol. 12, no. 23, Dec. 2021, doi: 10.23913/ride.v12i23.1103.
- [14]. W. Sandoval, "Análisis de Componentes Principales," *rstudio-pubs*.
- [15]. C. C. Aggarwal, *Neural Networks and Deep Learning*. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-319-94463-0.
- [16]. S. S. Alquegui, "Design and implementation of an indoor positioning system based on wireless embedded devices and Machine Learning," 2022. Accessed: May 28, 2023. [Online]. Available: <https://repositorio.unican.es/xmlui/bitstream/handle/10902/25230/441668.pdf?sequence=1>
- [17]. S. Weidman, *Deep learning from scratch*, First. O'Reilly, 2019.
- [18]. L. L. Lorente-Leyva, J. F. Pavón-Valencia, Y. Montero-Santos, I. D. Herrera-Granda, E. P. Herrera-Granda, and D. H. Peluffo-Ordóñez, "Artificial Neural Networks for Urban Water Demand Forecasting: A Case Study," *J Phys Conf Ser*, vol. 1284, no. 1, p. 12004, 2019, doi: 10.1088/1742-6596/1284/1/012004.
- [19]. L. Deng, M. Yang, H. Li, M. R. Lyu, and I. King, "Enhancing the reliability of out-of-distribution image detection in neural networks. ," Mar. 2021.
- [20]. A. Kaur and R. Saluja, "Handling Class Imbalance Problem Using Oversampling Techniques: A Review. In *Advances in Computing and Data Sciences* ," Springer, Singapore, pp. 9–19, 2020.
- [21]. Z. Zhang, S. Zohren, and S. Roberts, "DeepLOB: Deep Convolutional Neural Networks for Limit Order Books," vol. 6, Aug. 2020, doi: 10.1109/TSP.2019.2907260.
- [22]. C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019, doi: 10.1186/s40537-019-0197-0.
- [23]. S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," in *Procedia Computer Science*, Elsevier B.V., 2018, pp. 679–688. doi: 10.1016/j.procs.2018.05.069.
- [24]. A. E. Jácome Ortega, J. A. Caraguay Procel, E. P. Herrera-Granda, and I. D. Herrera Granda, "Confirmatory Factorial Analysis Applied on Teacher Evaluation Processes in Higher Education Institutions of Ecuador," in *Advances in Intelligent Systems and Computing*, Springer, 2020, pp. 157–170. doi: 10.1007/978-3-030-37221-7_14.

- [25]. E. P. Herrera-Granda, I. D. Herrera-Granda, L. L. Lorente-Leyva, P. D. Granda-Gudiño, and J. A. Caraguay-Procel, "Implementation of an Artificial Vision System and Monitoring of Human Objectives using a Quadricopter," *Revista Ibérica de Sistemas e Tecnologías de Informação*, vol. 2019, no. E19, pp. 198–211, 2019, [Online]. Available: <https://www.proquest.com/openview/ca355d43bc1ac0d236ca8da769393332/1?pq-origsite=gscholar&cbl=1006393>
- [26]. E. P. Herrera-Granda, J. C. Torres-Cantero, A. Rosales, and D. H. Peluffo-Ordóñez, "A Comparison of Monocular Visual SLAM and Visual Odometry Methods Applied to 3D Reconstruction," *Applied Sciences*, vol. 13, no. 15, p. 8837, Jul. 2023, doi: 10.3390/app13158837.
- [27]. E. P. Herrera-Granda, A. L. Díaz-Chimbo, P. Granda-Gudiño, M. R. Pusedá-Chulde, and I. D. García-Santillán, "Prototype of an Android mobile application for real time drowsiness detection and alertness applied in night driving," *Revista Ibérica de Sistemas e Tecnologías de Informação*, vol. 2020, no. E32, pp. 98–112, 2020, [Online]. Available: <https://www.proquest.com/openview/63deba1836c13b0fe0b01069ba7cda50/1?pq-origsite=gscholar&cbl=1006393>
- [28]. M. T. Hagan, H. B. Demuth, M. H. Beale, and O. De Jesús, *Neural Network Design 2nd Edition*, 2nd ed. 2014.
- [29]. E. M. Quindemil Torrijo and F. Rumbaut León, "Análisis de componentes principales para obtener indicadores reducidos de medición en la búsqueda de información," 2019. [Online]. Available: <https://orcid.org/0000-0003-3705-6297>
- [30]. M. I. C. Rachmatullah, J. Santoso, and K. Surendro, "Determining the number of hidden layer and hidden neuron of neural network for wind speed prediction," *PeerJ Comput Sci*, vol. 7, pp. 1–19, 2021, doi: 10.7717/PEERJ-CS.724.

Anexos

Las imágenes de las arquitecturas de las mejores configuraciones de redes neuronales se encuentran en: <https://goo.su/KMFox>