

Evaluación del Estado Emocional de los empleados de la empresa FenixCorp-ADS mediante el uso de Machine Learning

Evaluation of the Emotional State of the employees of the company FenixCorp-ADS using Machine Learning

Yomara Elizabeth Tello-Oña ¹
Universidad Técnica de Cotopaxi - Ecuador
yomara.tello8402@utc.edu.ec

Luis René Quisaguano-Collaguazo ²
Universidad Técnica de Cotopaxi - Ecuador
luis.quisaguano1@utc.edu.ec

doi.org/10.33386/593dp.2024.6.2695

V9-N6 (nov-dic) 2024, pp 192-203 | Recibido: 02 de agosto del 2024 - Aceptado: 17 de agosto del 2024 (2 ronda rev.)

1 ORCID: <http://orcid.org/0009-0000-9026-6950>

2 ORCID: <http://orcid.org/0000-0003-1345-0898>

Descargar para Mendeley y Zotero

RESUMEN

La Inteligencia Artificial (IA) tiene hoy en día un gran auge, no solo en los procesos industriales, sino también, en campos relacionadas con las actividades que contribuyen al desarrollo de las organizaciones y del conocimiento humano. El Aprendizaje Automático se encuentra contenida dentro del contexto computacional de la IA, dividida en tres grandes enfoques: Aprendizaje Automático Supervisado (AAS), Aprendizaje Automático No Supervisado y el Aprendizaje por Reforzamiento, que son técnicas y algoritmos capaces de “aprender y razonar” simulando al cerebro humano, y permite clasificar y predecir el comportamiento de los datos suministrado al modelo elegido, mediante la utilización de la biblioteca SciKit-Learn y otras herramientas de Python, lenguaje de programación muy utilizado en el análisis de datos y el análisis de sentimiento, otra herramienta para la extracción de información en base de opiniones subjetivas. El Estado Emocional (EE) es una reacción consciente e inconsciente como respuesta a un estímulo específico en un momento o situación que tiene los individuos. Con el fin de establecer una posible relación con el desempeño laboral de los empleados y la fuga de talento de la empresa FenixCorp-ADS, se elaboró un cuestionario vía Web y se eligieron los algoritmos de clasificación y predicción de aprendizaje automático supervisado y no supervisado que permitió elaborar un análisis comparativo para determinar el modelo más eficiente de los análisis de sentimientos efectuados, contribuyendo en su aplicación de forma más continua por las organizaciones empresariales a modo de su aprovechamiento para automatizar estos procesos.

Palabras claves: inteligencia artificial, análisis de datos, lenguaje de programación, actitud laboral.

ABSTRACT

Artificial Intelligence (AI) is currently booming, not only in industrial processes, but also in fields related to activities that contribute to the development of organizations and human knowledge. Machine Learning is contained within the computational context of AI, divided into three main approaches: Supervised Machine Learning (AAS), Unsupervised Machine Learning and Reinforcement Learning, which are techniques and algorithms capable of "learning and reasoning" by simulating the human brain, and allows classifying and predicting the behavior of the data supplied to the chosen model, through the use of the SciKit-Learn library and other Python tools, a programming language widely used in data analysis and sentiment analysis, another tool for extracting information based on subjective opinions. Emotional State (ES) is a conscious and unconscious reaction in response to a specific stimulus at a specific time or situation that individuals have. In order to establish a possible relationship with the job performance of employees and the talent drain of the company FenixCorp-ADS, a questionnaire was developed via Web and the classification and prediction algorithms of supervised and unsupervised machine learning were chosen, which allowed a comparative analysis to determine the most efficient model of sentiment analysis performed, contributing to its application in a more continuous way by business organizations in order to use it to automate these processes.

Keywords: artificial intelligence, data analysis, programming language, work attitude.

Introducción

La fuga de talentos por empleados insatisfechos o debido a un clima laboral estresante es un desafío significativo ante el cual las organizaciones deben enfrentarse en la actualidad. Estos parámetros pueden causar muchas dificultades a nivel empresarial y afectar considerablemente el Estado Emocional (EE) de los empleados, componente que forma parte del acontecer cotidiano de los individuos que se generan al interactuar con su entorno de manera constante en un momento determinado. Una alta competitividad y las exigencias propias de las empresas son propicios para crear un ambiente laboral estresante, jugando un rol importante en las tareas y desempeño laboral de los trabajadores. El desempeño laboral depende de ciertos factores (motivación, relaciones sociales, autorregulación, conciencia laboral, empatía, etc.) y no han tenido la debida importancia en las organizaciones, generando así, grandes retos en el desenvolvimiento laboral de los individuos (Aragón Zepeda, 2019).

El uso de plataformas tecnológicas como la IA, el Aprendizaje Automático (AA, también denominado Machine Learning) y el Análisis de Sentimiento (AS), permiten la identificación de patrones, comportamientos y/o expresivos de los individuos sujetos a estudios en un determinado momento de su EE dentro de un mundo de datos. Esto resulta evidente para la percepción humana, pero se tiene el reto de cuantificarlo con el fin de reducir significativamente el grado de incertidumbre e inexactitud generada por la manipulación de grandes volúmenes de datos a fin de evaluar su relación con el desempeño laboral.

El objetivo de este artículo es evaluar el EE en los empleados de la empresa FenixCorp-ADS, mediante las aplicaciones de pruebas y entrenamiento usando algoritmos de clasificación y predicción de ML como K-Means para AA no supervisado, Regresión Logística (RL) y Máquinas de Soportes Vectoriales (SVM) para el AA supervisado y no supervisado, contenidos en la biblioteca SciKit-Learn de Python y otras herramientas informáticas que

fueron aplicados en un AS a fin de diagnosticar una posible relación o influencia del EE de los trabajadores de la empresa con su desempeño laboral y la fuga de talentos, que conllevó a un estudio comparativo de los resultados obtenidos con la técnicas antes mencionadas.

El AA, a través del AS ofrece una solución innovadora y aborda la problemática sobre la fuga de talento, brindando herramientas que contribuyen en una evaluación objetiva y continua del EE de los empleados, en este caso específico se centraron en textos recabados mediante una aplicación Web propia en FenixCorp-ADS, empresa con más de 20 años de trayectoria dedicada al desarrollo de software contables y administrativos, con sucursales en Colombia, Ecuador y Perú, contando a su vez con 26 empleados y que ha sufrido una fuga de talento humano significativa en los últimos años, sin saberse a ciencia cierta los motivos o causas que lo originaron. Al utilizar los algoritmos de AS de Machine Learning (ML), se logró identificar y predecir patrones específicos asociados a los problemas emocionales de los empleados mediante el análisis de información relacionada con la fuga de talento.

Se revisó trabajos de investigación relacionados destacando la propuesta de Carceller Llorens, (2023) que en su tesis de maestría: Detección del EE usando técnicas de ML, elaboró un análisis de expresiones faciales de usuarios en tiempo real, utilizando bibliotecas Python, se usó los algoritmos Transfer Learning (TL) y Deep Learning (DP) evaluando las expresiones faciales en una red neural, lográndose mejores resultados con TL. DL no ofreció los resultados esperados.

Saura et al., (2018) aborda el AS en Twitter con Machine Learning, identificando el sentimiento sobre ofertas de #Black Friday, manifiesta que la red social Twitter es excelente para que usuarios y audiencia expresen sus sentimientos, opiniones y comentarios en una temática específica, concluyendo que luego del análisis de los datos, arrojaron los resultados siguientes: 60,2 % neutros, 32,1% positivos y 7,7 % negativos, y de acuerdo a ello pueden servir

de guía en empresas relacionadas con marketing a través de las redes sociales. Por otra parte Alcalde Chulilla, (2021) en su tesis de grado AS de textos basado en opiniones de películas usando ML, utilizó técnicas de Procesamiento de Lenguaje Natural (NLP), creando un conjunto de algoritmos para su análisis a partir de la base de datos de Internet Movie Data Base (IMDB), utilizando los algoritmos TL, RL y Naïve Bayes Multinomial, obteniendo resultados favorables de hasta un 90 %, donde TL fue el más destacado en los resultados obtenidos.

El desafío para analizar de manera precisa y no intrusiva del EE a los trabajadores de la empresa, no es recurrir a la evaluación tradicional basadas en encuestas, cuestionarios o entrevistas de forma física, que a menudo suelen disponer de recursos y tiempo adicionales, además de ser subjetivas porque no permiten obtener resultados eficientes por la complejidad y variabilidad de las emociones humanas en un ambiente laboral dinámico, sino más bien es la de enfocar la solución del problema a las plataformas tecnológicas que ofrecen el AA mediante en el uso algoritmos de AS para gestionar en forma eficiente el bienestar y el rendimiento laboral de las organizaciones y que pueden afectar el clima laboral y la cultura organizacional de la empresa FenixCorp-ADS, quienes han brindado su total apoyo para el desarrollo de esta investigación.

Marco conceptual

La inteligencia artificial y el aprendizaje automático

La IA abarca la combinación de diferentes plataformas tecnológicas que al aplicarse en diversas maneras constituyen sistemas capaces de percibir, comprender e interactuar con su entorno, basándose fundamentalmente en la capacidad de “aprender” a partir de datos proporcionados por los humanos. La IA, se conceptualiza como una ciencia y/o un conjunto de tecnologías computacionales que se centran en sistemas y programas diseñados para emular algunas de las facultades intelectuales humanas, tales como el pensamiento, el razonamiento, el aprendizaje y la toma de decisiones autónomas,

que pueden ser consideradas como inteligentes (Andrade Muñoz, (2023) y (Olivas et al., 2023).

Dentro de las áreas que contienen la IA, se encuentra el AA, disciplina con aplicaciones y algoritmos capaces de aprender y adaptarse para la resolución de problemas de clasificación y predicción sin la imperiosa necesidad de ser programados explícitamente, logrando a su vez que a partir de los datos puedan inferir para la resolución de problemas (Velásquez Sánchez, 2022). El aprendizaje automático se clasifica de acuerdo con tres enfoques diferentes, según Bobadilla, (2020):

- **Aprendizaje supervisado:** Algoritmos entrenados, que aprende utilizando datos categorizados en clases (etiquetas), y permiten obtener predicciones más precisas en un nuevo conjunto de datos. En esta categoría se encuentran dos tipos: los de clasificación en que los algoritmos encuentran patrones dentro de los datos originales, los clasifica en grupos para luego compararlos y alojarlos en subgrupos y predecir su comportamiento, devolviendo valores discretos y los de regresión, modelan la relación entre variables, devolviendo un valor continuo entre 0 y 1, utilizando el margen de error y prediciendo la asociación de la variable independiente con la dependiente.

- **Aprendizaje no supervisado:** Estos modelos no están categorizados en clases o etiquetas y tratan de agrupar los datos originales identificando sus características y similitudes comunes. Evalúan si existe alguna correlación o vínculo, reorganizando los datos de forma ordenada a medida que analizan nuevos datos. Los hay de dos tipos: por agrupamiento (clustering) y por reducción de la dimensionalidad (preprocesamiento y visualización de los datos).

- **Aprendizaje por refuerzo:** Algoritmos que se entrenan a través de la interacción constante mediante pruebas de ensayo y error en un entorno específico a partir de experiencias y datos previos. El algoritmo intenta explorar varias opciones y perspectivas, monitoreando y evaluando cada salida para determinar cuál es la ideal, y comienza a cambiar su enfoque

de la situación para alcanzar el mejor resultado posible.

La característica más importante del AA, son los datos; estos deben ser suficientes y de calidad para que el modelo a utilizar pueda generalizarlos, para ello, es necesario realizar previamente un análisis exploratorio, mediante la construcción de un conjunto de datos con la finalidad de lograr la unicidad y reducir la incertidumbre e inexactitud en los resultados. Esto ha llevado a La IA y el AA, a tener un auge importantísimo hoy en día, incluso ha generado debates y discusiones sobre sus implicaciones éticas y sociales, como la disminución de las fuentes de trabajo y la privacidad de los datos y también en algunos aspectos legales como las violaciones a la propiedad intelectual (Andrade Muñoz, 2023).

En la figura1, se aprecia los diferentes algoritmos más utilizados en Python de acuerdo con su enfoque y tipos:

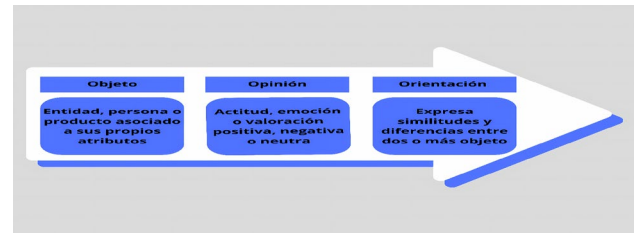
Figura 1
Tipos de algoritmos, según su enfoque



Análisis de sentimiento

Es un área de estudio computacional para evaluar las emociones, actitudes y opiniones de los individuos sujetos a análisis a partir de etiquetas o datos textuales. El AS detecta las expresiones positiva, negativa o neutra (etiquetas) sobre una temática específica a través de tres componentes: un objeto, opinión y la orientación de la opinión de un elemento textual (Rosenbrock et al., 2021), como se indica en Figura 2.

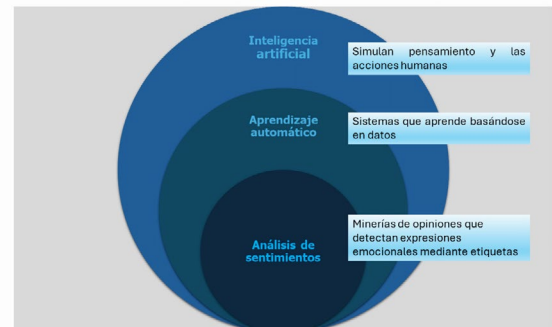
Figura 2
Etiquetas para análisis de sentimiento



Nota: Las etiquetas están contenidas en el enfoque de AAS, tienen una ponderación de confianza entre 0 y 1 para todo el texto del documento analizado. Basado en (Rosenbrock et al., 2021)

Como contexto computacional el AS, el AAS y la IA, son plataformas tecnológicas contenidas una dentro de la otra que contribuyen de manera significativa a inferir en la resolución de problemas de predicción y clasificación de textos. La siguiente figura 3 se detalla el contexto de estas disciplinas.

Figura 3
Contexto computacional de la inteligencia artificial



El estado emocional y la fuga de talento

El EE es una reacción de los individuos como respuesta al interactuar con su entorno y que forma parte de la vida diaria de los individuos. Surgen de la valoración consiente e inconsciente de un estímulo específico en un momento o situación determinada como respuesta a los objetivos de los individuos y de acuerdo con Blanco Canales, (2019) pueden ser causadas por los factores siguientes:

- **Fisiológicos:** emociones donde intervienen los sistemas nervioso central y el endocrino, así como la homeostasis emocional.

- **Conductuales:** todo lo relacionado a las actividades del movimiento corporal, expresiones faciales, la voz, entre otros, donde interviene las expresiones corporales y verbales.

- **Reacciones psicológicas:** emociones de los individuos ante los cambios sugestivos en determinados momentos o situaciones.

Aunque el EE está muy relacionada a los sentimientos, estos son los que reflejan la autopercepción a raíz de una emoción específica, es la experiencia vivida por la reacción emocional.

El EE se encuentra muy asociado en el ámbito laboral debido a su ambiente estresante por la competitividad y la falta de motivación y puede manifestarse con cuadros de ansiedad, depresión y síntomas de agotamiento físico y emocional en los trabajadores. A estos factores de estrés laboral lo han denominado: Síndrome de Burnout, “como un estado psicológico relacionado con el trabajo negativo y persistente, presente en empleados de aspecto normal que presentan síntomas de incompetencia, malestar, desmotivación y disfuncionalidad laboral” (Lovo, 2020), que puede repercutir en la fuga de talento y de aumentar el ambiente estresante en las organizaciones empresariales.

En base a la investigación realizada, las técnicas de ML tienen un ámbito de acción muy variada, sus algoritmos han sido aplicados para análisis de datos en los sectores de ciencias de la salud, marketing digital, la industria manufacturera y la psicología. Sin embargo, en esta última área ha estado limitada a la psicología cognitiva, es decir, la solución de problemas y razonamiento a través de los datos suministrados al sistema, que les permiten “aprender y razonar” la toma de decisiones, no así, en lo correspondiente al AS con respecto a la fuga de talento, valioso activo que tienen hoy el día las organizaciones permitiéndoles identificar patrones de conducta o escenarios en los EE de los trabajadores para una posible retirada. La finalidad de este artículo

es suministrar conocimiento científico que permitan evaluar con precisión de estas conducta o posibles escenarios a fin de conocer las causas que la generan.

Metodología

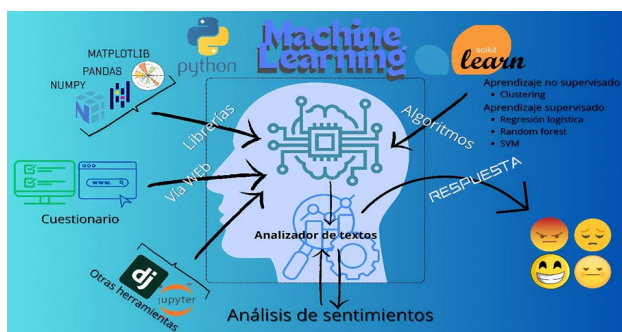
No existe una metodología concreta a seguir para los proyectos de aplicaciones de técnicas y algoritmos de ML. En este artículo se utilizó la metodología tradicional con técnicas del método científico, y de acuerdo con el nivel de conocimientos, la investigación es del tipo descriptiva, debido a la recolección de datos a partir de la evaluación e interpretación correcta de los datos de una población (Tamayo y Tamayo, 2003, p. 40), para luego desarrollar procedimientos y técnicas algorítmicas de ML y emplearlas en la clasificación del análisis de datos, siguiendo las pautas y/o procedimientos del método inductivo, con el fin de establecer la relación de las características en que se presentan en sucesos determinados. El enfoque mixto (cualitativo-cuantitativo) permite la evaluación de la predicción de las variables continuas y discretas. Este artículo pretende ofrecer conocimiento científico sobre el AA y el AS de acuerdo con los resultados analíticos obtenidos para ofrecer soluciones personalizadas y oportunas a fin de mejorar el bienestar y el rendimiento laboral de los empleados de la empresa y otras PYME.

A través de un AS como herramienta del ML, se puede esperar resultados prometedores en la evaluación y detección de los sentimientos de los empleados de la empresa para determinar la existencia relacional entre EE y el desempeño laboral, ahorrándose tiempo y recursos, así como la reducción de la incertidumbre e inexactitud generadas por las variables en estudio. Este artículo tiene como objeto de estudio El estado emocional de los empleados de FenixCorp-ADS: variable dependiente y mediante la aplicación de técnicas de un AS con ML, se pretende determinar la hipótesis de investigación (H_1): el Síndrome de Burnout y el EE de los trabajadores de la empresa FenixCorp-ADS influye con la fuga de talento y el desempeño laboral y la hipótesis nula (H_0): el Síndrome de Burnout y el EE de los trabajadores

de la empresa FenixCorp-ADS no influye con la fuga de talento y el desempeño laboral

El campo de acción son los Algoritmos de ML utilizados para el análisis de datos como la variable independiente. La figura 4 detalla de forma esquemática la metodología empleada para todas las herramientas y librerías utilizadas de los modelos de algoritmos de SciKit-Learn y Python.

Figura 4
Esquema metodológico de las aplicaciones utilizadas



Población y muestra

La población para responder el cuestionario que fue considerada (Tabla 1) en su totalidad de los 26 empleados con lo que cuenta de la empresa, dado a que es un número manejable para la investigación no se realizó un proceso de muestreo, debido a esto no se tomó muestra de la población detallada en la en tabla 1:

Tabla 1
Dato poblacional de empleados de la empresa FenixCorp-ADS

| Género | Cantidad |
|---------|----------|
| Hombres | 6 |
| Mujeres | 20 |
| Total | 26 |

Materiales y métodos utilizados

Se sistematizo un cuestionario digital mediante una aplicación web desarrollada con el framework Django para la recolección de datos por parte de los empleados de la empresa. Este cuestionario se basó en preguntas tomadas de

un test del Instituto Nacional de Estadísticas, Geografía e Informática (INEGI) de México, cuya finalidad es la construcción de un conjunto de datos de los EE a partir de las respuestas suministradas por los trabajadores de la empresa FenixCorp-ADS, en el lapso de un mes.

Para la evaluación del EE de los empleados a ser objetos de estudio, se emplearon herramientas y aplicaciones informáticas siguiendo el siguiente orden:

- Construcción de un conjunto de datos a partir de las respuestas formuladas en un cuestionario enviado a los empleados vía web. Para ello, este cuestionario se diseñó utilizando el framework Django y el lenguaje JavaScript para conexión del dominio y servidor alquilado para tal fin.
- Se utilizaron las librerías Matplotlib, Numpy y Pandas de Python para preparación, limpieza y normalización del conjunto de datos que contienen las variables de los EE de los empleados de la empresa.
- Se emplearon técnicas y algoritmos ML contenidas en la librería SciKit-Learn de Python para la clasificación y predicción de las variables, así como el entrenamiento de los modelos de datos obtenidos.
- Se evaluaron los resultados de acuerdo con los gráficos y matrices de confusión generadas por los algoritmos.

Herramientas utilizadas para desarrollar modelos de Machine Learning

Existen varias herramientas y lenguajes de programación con sus librerías y aplicaciones para desarrollar modelos de ML de clasificación y predicción en el análisis de datos. En la elaboración de este artículo, solo se tomó en cuenta Python v3.11 y sus aplicaciones compatibles y JavaScript, para el diseño de un sitio Web.

Python

Lenguaje de programación de código abierto y de propósito general e interpretativo, es decir no se compilan directamente, es muy utilizado por desarrolladores de sitios Web y científicos de datos en diferentes plataformas: Windows, Linux, MacOS. También contiene una gran variedad de bibliotecas que aumenta sus funcionalidades. Su librería principal SciKit-Learn posee los algoritmos principales para el análisis de datos de AA. La tabla 2 refleja las librerías y algoritmos para los diferentes modelos de ML utilizados.

Tabla 2
Software y aplicaciones utilizadas en Python

| Tipo | Nombre | Versión | Función |
|-----------|---------------------|---------|-----------------------------------------------------------------------------------|
| Librería | SciKit-Learn | 1.24.1 | Clasificación y predicción: K-Means, SVM, Regresión logística, polinomial, lineal |
| Librería | Pandas | 1.5.0 | Desarrollo de dataframe |
| Librería | NumPy | 1.26.0 | Diseño de vectores y matrices |
| Librería | Matplotlib, Seaborn | 3.8.2 | Diseño de gráficos estadísticos |
| Framework | Django | 5.6.0 | Diseño de sitio Web |
| Librería | LSTM | 2.0 | Creación de redes neurales recurrentes con Tensorflow y Keras |

Hardware

Para la evaluación del estado emocional de los empleados de la empresa FenixCorp-ADS se utilizó un equipo cuyas características se describen en Tabla 3.

Tabla 3
Características del computador utilizado

| | |
|-------------------|----------------------------------|
| Característica | Laptop |
| Marca | Dell |
| Procesador | Intel Core i5-1135G7 11Gen. |
| Memoria RAM | 12 Gb |
| Almacenamiento | 500 Gb |
| Sistema operativo | Linux Ubuntu Noble Numbat, v24.0 |

Para evaluar los resultados obtenidos por los modelos aplicados, se definieron las métricas más representativas de la librería SciKit-Learn, detalladas en las tablas 4 y 5.

Tabla 4
Métricas de regresión aplicadas en las pruebas realizadas

| Regresión | | |
|-------------------------------------------------|---------------------------------------|----------------------------------------------------|
| Métrica | Método o clase en Python | Función |
| Raíz del valor medio cuadrado (RMSE) | sklearn.metrics.mean_squared_error() | Raíz cuadrada del error medio cuadrado |
| Error medio cuadrado (MSE) | sklearn.metrics.mean_squared_error() | Valor cercano a la línea de ajuste de la regresión |
| Coefficiente de determinación (R ²) | sklearn.metrics.r2_score() | Variación entre el valor real y el de predicción |
| Error medio absoluto (MAE) | sklearn.metrics.mean_absolute_error() | Valor absoluto del error medio cuadrado |

Tabla 5
Métricas de clasificación

| Clasificación | | |
|--------------------------|---------------------------------------------------------------------|-----------------------------------------------------------|
| Métrica | Método o clase en Python | Función |
| Matriz de confusión | sklearn.metrics.confusion_matrix() | Evalúa la precisión mediante una matriz |
| Exactitud | sklearn.metrics.accuracy_score() | Evalúa la puntuación de la precisión del modelo |
| Precisión-Recall | sklearn.metrics.recall_score() sklearn.metrics.precision_score() | Evalúa la calidad del clasificador |
| Reporte de clasificación | sklearn.metrics.classification_report() | Elabora un informe comparativo de las métricas utilizadas |
| Medición F | sklearn.metrics.f1_score() | Valor medio de la precisión |

Resultados

Los resultados obtenidos de acuerdo con las diferentes pruebas realizadas en el estudio brindaron información con base a las métricas contenidas en la librería SciKit-Learn de Python y están detallados en los siguientes apartados.

Recopilación de los datos

Las respuestas al cuestionario fueron suministradas por los empleados de la empresa FenixCorp-ADS, mediante la interacción con un cuestionario web, elaborado para tal fin. En la figura 5 se observa una pantalla desarrollada para las respuestas.

Figura 5
Pantalla del cuestionario web elaborado con Django para la recopilación de datos

Tema: Evaluación del Estado Emocional de los empleados de la empresa FenixCorp-ADS mediante el uso de Machine Learning

1. Nada de acuerdo 2. Algo de acuerdo 3. Bastante de acuerdo 4. Muy de acuerdo 5. Totalmente de acuerdo

1. ¿Quiero prestar más atención a mis emociones afectan mi estado de ánimo en el trabajo?

2. ¿Normalmente dedico tiempo a pensar en mis emociones?

3. ¿Presta mucha atención a cómo me siento en mi trabajo?

4. ¿Fanzo en mi estado de ánimo constantemente?

5. ¿Frecuentemente puedo expresar mis sentimientos en mi lugar de trabajo?

6. ¿A menudo me doy cuenta de mis emociones en diferentes situaciones?

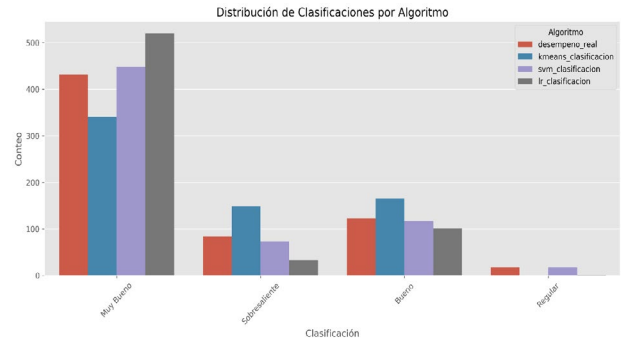
7. ¿Siempre puedo decir cómo me siento?

8. ¿Tengo claro mis sentimientos?

9. ¿Me presiono por tener un buen estado de ánimo?

10. ¿Alguna vez siento mal, procuro pensar en cosas agradables?

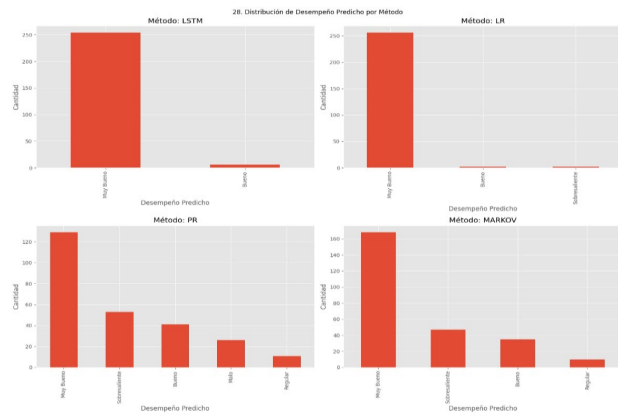
Figura 7
Distribución de la variable de desempeño laboral por cada algoritmo de clasificación utilizado



Preparación y visualización de los datos

Se realizó un análisis exploratorio previo al conjunto de datos objeto a evaluación, esto incluye la normalización y limpieza de los datos a través de las herramientas de Python para tal fin, que incluye gráficos e histogramas, como se puede observar en las figuras 6 y 7.

Figura 6
Distribución de los datos por modelo



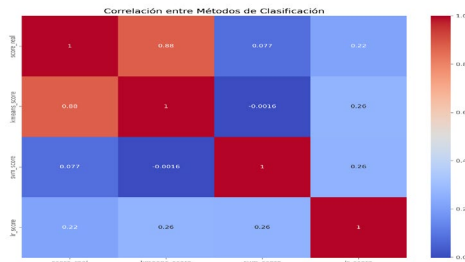
La matriz de correlación, también llamada matriz de correlación de Pearson, refleja los grados de conexión de los coeficientes entre diferentes variables o parámetros del conjunto de dato elaborado. Cada celda muestra el grado de correlación entre los parámetros. Al detallarse a nivel visual aporta información sobre la linealidad de los datos, pues en ella se comparan todas las posibles combinaciones del conjunto de datos, permitiendo además, de la visualización se pueden identificar patrones en los datos. Se contemplan además dos tipos de correlación:

Positiva: fuerte relación de la variable X con la variable Y (en ambos con valores altos).

Negativa: fuerte relación de la variable X con la variable Y (con valores altos y bajos).

En la matriz de correlación creada entre métodos de clasificación desde el conjunto de datos, se puede visualizar de forma directa que hay solo dos variables con un alto nivel de correlación positivo (K-Means score y score real). Se observa también que las variables contenidas en celdas de color azul no se relacionan con casi ninguno otro parámetro. En la figura 8, se puede visualizar con mayor detalle las diferentes celdas en que están compuesta la matriz de correlación.

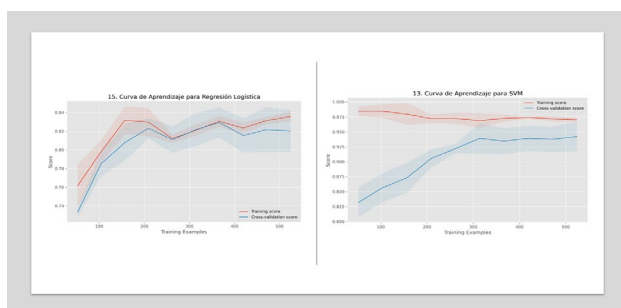
Figura 8
Matriz de correlación de Pearson



Entrenamiento y validación de los algoritmos

Se dividió en dos el conjunto de datos, uno para las pruebas de entrenamiento del algoritmo utilizado y pueda aprender de los diferentes parámetros contenidos en los documentos y el otro para pruebas que midan el rendimiento del modelo resultante permitiendo clasificarlos en nuevos documentos: sobresaliente, muy bueno, bueno y regular. Esto tiene la finalidad que los algoritmos puedan suministrar información valiosa de los datos agregados para luego poder realizar las predicciones. Se aplicaron los modelos: SVM y RL, ambos algoritmos de AAS fueron utilizados tanto para procesamiento de clasificación como de regresión y K-Means solo para clasificación. En la figura 9, se puede observar con detenimiento las diferentes etapas de entrenamiento y validación cruzada para los algoritmos RL y SVM.

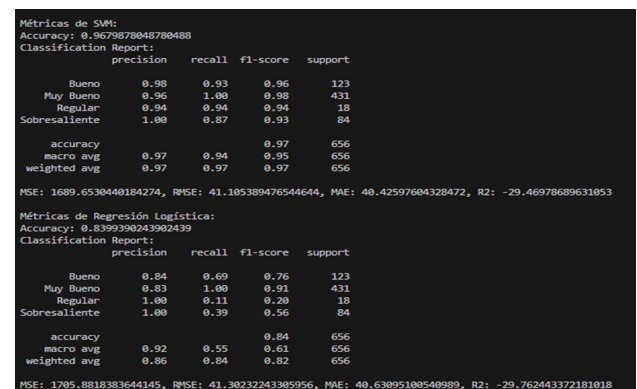
Figura 9
Curva de entrenamiento y validación cruzada para RL y SVM



Asimismo, se evaluó el desempeño de los algoritmos entrenados por medio de la clase reporte de clasificación de Python, se elaboró un reporte resumido mostrando las métricas utilizadas de clasificación con base a precisión,

recall, f1-score y support, así como la utilización de las principales métricas de regresión: MAE, RMSE, MAE y R². La figura 10, muestra el reporte de clasificación con cada una de las métricas utilizadas con sus respectivos valores obtenidos, por medio del entrenamiento y la validación.

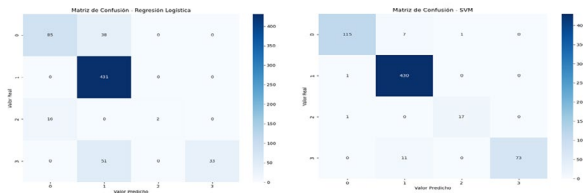
Figura 10
Reporte de clasificación con las métricas resultantes



Nota: Se puede observar que las métricas de clasificación obtenidas tienen valores mayores a 0.8, lo que hace que los modelos seleccionados sean muy confiables para el tratamiento de los datos.

Otra manera de medir el rendimiento de los modelos utilizados es a través de la matriz de confusión (ver figura 11) , estructurada en la manera siguiente: en la coordenada X se encuentran los valores predcibles con respecto a los valores reales (coordenada Y), además de observar dentro de esta matriz que los modelos de RL y SVM contiene valores confiables, debido a que la mayoría de los valores predicho han sido acertados correctamente (valores verdaderos positivos y verdaderos negativos), celdas de color azul más oscuros

Figura 11
Matriz de confusión para RL y SVM

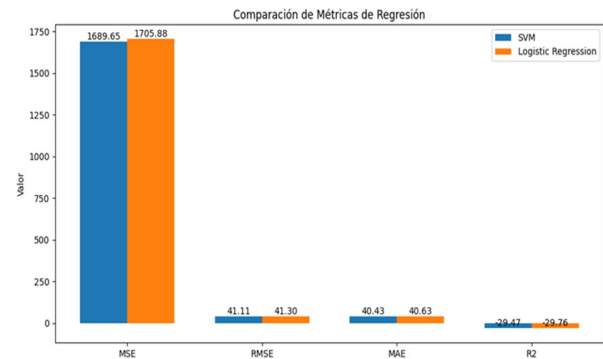


Con respecto a los resultados obtenidos sobre las métricas para regresión, también podemos visualizarla a través de la figura 10, y contenidas en la tabla 6. La figura 12, se aprecia una relación comparativa entre estas métricas de regresión y sus valores están muy cercanos a cero lo que las hacen muy confiables como validación de estos modelos. Sin embargo también se observa que la métrica R^2 , no es adecuada para el tratamiento de los datos, debido a que sus valores están muy por debajo de cero, dando un valor negativo.

Tabla 6
Resultados de las métricas para la validación de los modelos de regresión

| Métricas de Regresión | | |
|-----------------------------------------|--------------------------|---------------------------------------|
| Métrica | Regresión logística (RL) | Máquinas de Soporte Vectoriales (SVM) |
| Raíz del valor medio cuadrado (RMSE) | 41,30 | 41,30 |
| Error medio cuadrado (MSE) | 1.705,88 | 1.689,65 |
| Coefficiente de determinación (R^2) | -29,45 | -29,76 |
| Error medio absoluto (MAE) | 40,43 | 40,63 |

Figura 12
Comparativa de las métricas de regresión para RL y SVM



Conclusiones

Se utilizaron seis modelos de AAS, bajo dos enfoques: clasificación (RL y SVM) y regresión (RL, SVM, LSTM,) y uno no supervisado de agrupamiento: K-Means, con el objetivo de utilizar estas técnicas de ML para entrenar diferentes modelos que permitan predecir una relación o influencia entre el EE y el desempeño laboral de los empleados de la empresa FenixCorp-ADS, siguiendo las pautas establecidas en la metodología tradicional en cuanto al levantamiento de la información a partir de en otros estudios, lecturas y trabajos relacionados con el tema investigado.

Con la aplicación los algoritmos mencionados se han obtenidos resultados satisfactorios por el rendimiento de las métricas utilizadas tanto de clasificación y como de regresión con valores cercanos a los referenciados como óptimos, de acuerdo con el algoritmo utilizado. El modelo SVM ofreció los mejores resultados, por tanto, basándose en los indicadores de las métricas obtenidas los experimentos en la aplicación de dichos algoritmos resultaron positivos y eficiente, muy cercano a los valores reales.

Hay que seguir proponiendo trabajos futuros que sigan esta línea de investigación por las ventajas que ofrece el ML y sus aplicaciones en el análisis de sentimiento para contribuir en la toma de decisiones en cuanto a la fuga de

talentos, que se ha convertido el gran problema para las organizaciones empresariales.

Referencias bibliográficas

- Alcalde Chulilla, J. (2021). *Análisis de sentimiento de textos basado en opiniones de películas usando algoritmos de aprendizaje computacional* [Tesis de grado, Universidad Oberta de Catalunya]. <https://openaccess.uoc.edu/bitstream/10609/132328/7/jchulillaTFG0621memoria.pdf>
- Andrade Muñoz, J. (2023). Entendiendo el poder de la Inteligencia Artificial. *TEPEXI, Boletín Científico de la Escuela Superior del Rio*, 10(20), 4. <https://repository.uaeh.edu.mx/revistas/index.php/tepexi/issue/archive>
- Aragón Zepeda, K. I. (2019). Inteligencia emocional y su relación en el desempeño laboral. *Revista Naturaleza, Sociedad y Ambiente*, 6(1), 57-67. <https://doi.org/10.37533/cunsurori.v6i1.41>
- Blanco Canales, A. (2019). La emoción y sus componentes. *Grupo LEIDE*, 7. <https://grupoleide.com/wp-content/uploads/2019/09/Ana-Blanco-y-Nati-Hern%C3%A1ndez-C%C3%B3mo-sentimos-la-L2.pdf>
- Bobadilla, J. (2020). *Machine learning y deep learning. Usando Python*. Ediciones de la U. https://api.pageplace.de/preview/DT0400.9789587921465_A41974869/preview-9789587921465_A41974869.pdf
- Carceller Llorens, F. (2023). *Detección de estados de ánimo usando técnicas de machine learning* [Tesis Magister, Universidad Politécnica de Valencia]. <http://polipapers.upv.es/index.php/IA/article/view/3293>
- Lovo, J. (2020). Síndrome de burnout: Un problema moderno. *Entorno*, 70, 110-120. <https://doi.org/10.5377/entorno.v0i70.10371>
- Olivas, J. Á., Montoro, A., & Lorenzo, A. (2023). *Informe OBS Inteligencia Artificial* (p. 38). OBS Business School. <https://marketing.onlinebschool.es/Prensa/Informes/Informe%20OBS%20Inteligencia%20Artificial%202023.pdf>
- Rosenbrock, G., Trossero, S., & Pascal, A. (2021). Técnicas de Análisis de Sentimientos Aplicadas a la Valoración de Opiniones en el Lenguaje Español. *CACIC 2021 UNSa*, 291-300. https://www.researchgate.net/publication/355887680_Tecnicas_de_Analisis_de_Sentimientos_Aplicadas_a_la_Valoracion_de_Opiniones_en_el_Lenguaje_Espanol
- Saura, J. R., Reyes Menéndez, A., & Palos Sánchez, P. (2018). Un análisis de sentimiento en Twitter con Machine Learning: Identificando el sentimiento sobre oferta de #Black Friday. *Revista Espacios*, 39(42), 16. <https://www.revistaespacios.com/a18v39n42/a18v39n42p16.pdf>
- Tamayo y Tamayo, M. (2003). *El proceso de la investigación científica* (4.ª ed.). Limusa, S.A.
- Velásquez Sánchez, E. (2022). *Estado del Arte de Machine Learning y su Aplicación en el Experimento LHCB* [Estado del arte de maestría en matemáticas, Universidad Nacional Autónoma de Honduras]. <https://mm.unah.edu.hn/dmsdocument/13661-estado-del-arte-de-machine-learning-y-su-aplicacion-en-el-experimento-lhcb-pdf>