**INFORME DE GUÍA PRÁCTICA**

1. **PORTADA**

Tema: GUIA APE 3

Unidad de Organización Curricular: PROFESIONAL

Nivel y Paralelo: Nivel - Paralelo

Alumnos participantes: Castro Iza Bryan Vladimir

Tipan Nuñez René Sebastian

Asignatura: Inteligencia de negocios

Docente: Ing. Edison Álvarez

1. **INFORME DE GUÍA PRÁCTICA**
2. **PP**
3. **YY**
   1. **Objetivo**

Construir un modelo de Regresión lineal

* 1. **Modalidad**

Presencial

* 1. **Tiempo de duración**

**Presenciales:** 10

**No presenciales:** 0

* 1. **Instrucciones**

Descargue el archivo dataset\_Facebook.csv

Cárguelo en Python, librería Pandas

Modifique las etiquetas si es necesario para un mejor análisis.

Divida los datos en entrenamiento y prueba

Aplique el modelo de Regresión Lineal Múltiple a los datos de entrenamiento

Compruebe el modelo con los datos de prueba

* 1. **Listado de equipos, materiales y recursos**
* Inteligencia artificial, TAC
* Computador
* Python

TAC (Tecnologías para el Aprendizaje y Conocimiento) empleados en la guía práctica:

Plataformas educativas

Simuladores y laboratorios virtuales

Aplicaciones educativas

Recursos audiovisuales

Gamificación

Inteligencia Artificial

Otros (Especifique): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

* 1. **Actividades por desarrollar**

**Descarga y preparación del dataset**

Se descargó el archivo dataset\_Facebook.csv proporcionado en la guía.

Se realizó la carga del conjunto de datos en el entorno Python utilizando la librería Pandas, mediante el comando:



Se eliminaron columnas que se consideraron irrelevantes para el análisis.

A screenshot of a computer

AI-generated content may be incorrect.

Se verificaron los nombres de las columnas y se modificaron las etiquetas con nombres más descriptivos, eliminando espacios y caracteres especiales para facilitar el análisis posterior.

A screenshot of a computer

AI-generated content may be incorrect.

Análisis exploratorio de datos (EDA)

Se evaluaron los valores faltantes y se eliminaron registros incompletos.



Se analizaron medidas estadísticas (media, desviación estándar, valores máximos y mínimos) y correlaciones entre las variables.

Se utilizaron gráficos de dispersión y mapas de calor (heatmaps) para observar la relación entre las variables predictoras y la variable dependiente.

Selección de variables relevantes

A partir del análisis de correlación, se seleccionaron las variables con mayor relación con la variable dependiente (por ejemplo: alcance o total\_interacciones).

División del conjunto de datos

Se dividió el dataset en dos subconjuntos:

Entrenamiento: 80% de los datos

Prueba: 20% de los datos

Implementación en Python:

A screenshot of a computer

AI-generated content may be incorrect.

Construcción del modelo de Regresión Lineal Múltiple

Se entrenó el modelo con los datos de entrenamiento, obteniendo los coeficientes y el intercepto de la ecuación lineal.

Evaluación del modelo

Se aplicó el modelo a los datos de prueba y se calcularon las métricas de desempeño:

A screenshot of a computer

AI-generated content may be incorrect.

R² (Coeficiente de determinación)

MSE (Error Cuadrático Medio)

MAE (Error Absoluto Medio)

A screenshot of a computer

AI-generated content may be incorrect.

Se obtuvo un R² de 1.00, evidenciando un buen ajuste del modelo.

Visualización de resultados

Se graficó la comparación entre los valores reales y los predichos para observar la precisión del modelo:

A screen shot of a computer

AI-generated content may be incorrect.

Se interpretaron los coeficientes obtenidos, identificando qué variables tienen mayor peso sobre la variable dependiente.

Finalmente, se elaboró el presente informe con los resultados, conclusiones y recomendaciones del proceso completo.

* 1. **Resultados obtenidos**

Después de cargar y analizar el conjunto de datos dataset\_Facebook.csv mediante la librería **Pandas**, se realizó una limpieza de datos, ajustando etiquetas y eliminando valores nulos para garantizar la calidad del modelo. Posteriormente, se dividió la información en dos subconjuntos: **entrenamiento (80%)** y **prueba (20%)**.

Con el conjunto de entrenamiento se aplicó un modelo de **Regresión Lineal Múltiple** utilizando la librería **Scikit-learn (LinearRegression)**, considerando como variables independientes los factores de interacción de las publicaciones (por ejemplo: número de comentarios, compartidos y tipo de publicación) y como variable dependiente el **alcance total** o **nivel de interacción**.

El modelo se ajustó correctamente, obteniendo un **coeficiente de determinación (R²) = 1.00**, lo que indica que el 100% de la variabilidad de la variable dependiente puede explicarse por las variables predictoras.

El **error cuadrático medio (MSE)** fue bajo, reflejando una buena capacidad de generalización del modelo sobre los datos de prueba.

Gráficamente, la línea de regresión mostró una tendencia positiva entre las variables predictoras y el alcance de las publicaciones, lo que sugiere una relación directa entre la actividad del usuario y el rendimiento del contenido publicado.

* 1. **Habilidades blandas empleadas en la práctica**

Liderazgo

Trabajo en equipo

Comunicación asertiva

La empatía

Pensamiento crítico

Flexibilidad

La resolución de conflictos

Adaptabilidad

Responsabilidad

* 1. **Conclusiones**

El modelo de Regresión Lineal Múltiple permitió identificar las variables que tienen mayor influencia en el alcance de las publicaciones, demostrando que factores como el número de comentarios, reacciones y compartidos son determinantes en el rendimiento del contenido.

El valor obtenido de R² = 1.00 evidencia que el modelo posee una buena capacidad explicativa, lo cual confirma la pertinencia del uso de la regresión lineal para analizar comportamientos en redes sociales.

Durante el proceso de análisis se comprobó que no todas las variables independientes aportan significativamente al modelo, por lo que la selección de características es esencial para mejorar la precisión y evitar sobreajuste.

Se verificó la importancia de aplicar correctamente las fases del proceso de minería de datos: limpieza, transformación, entrenamiento y evaluación, las cuales garantizan resultados confiables y replicables.

La práctica contribuyó al fortalecimiento de habilidades analíticas y al entendimiento de cómo los modelos predictivos pueden ser utilizados para la toma de decisiones en entornos empresariales o de marketing digital.

* 1. **Recomendaciones**

Antes de aplicar modelos predictivos, se recomienda realizar una evaluación exploratoria detallada de las variables para eliminar redundancias y detectar posibles valores atípicos que afecten el desempeño del modelo.

Se sugiere probar otros algoritmos de regresión (como Ridge o Lasso) para comparar resultados y mejorar la capacidad de generalización del modelo.

Mantener un equilibrio entre la cantidad de variables y la simplicidad del modelo, ya que incluir demasiadas puede generar ruido y sobreajuste.

Documentar cada fase del proceso analítico para asegurar la reproducibilidad de los resultados y facilitar futuras mejoras.

Finalmente, se recomienda profundizar en la interpretación de los coeficientes de regresión, de modo que las predicciones puedan transformarse en estrategias accionables dentro del contexto de inteligencia de negocios.

* 1. **Referencias bibliográficas**

[1] J. Jironés, Minería de datos: modelos y algoritmos, 1ª ed., Editorial UOC, Barcelona, España, 2017.

[2] M. A. Farooq y B. A. Masood, SQL Server 2014 Development Essentials, 1st ed., Packt Publishing, Birmingham, Reino Unido, 2014.

[3] C. Pérez López, Minería de datos: técnicas y herramientas, 1ª ed., Thomson, Madrid, España, 2007.

[4] M. Pérez Marqués, Minería de datos a través de ejemplos, 1ª ed., Alfaomega, México D.F., México, 2015.

[5] S. Raschka y V. Mirjalili, Python Machine Learning, 3rd ed., Packt Publishing, Birmingham, Reino Unido, 2019.

[6] W. McKinney, Python for Data Analysis, 2nd ed., O’Reilly Media, Sebastopol, EE. UU., 2017.

* 1. **Anexos**