

# Informe del Proyecto: Sistema de Recomendación de Juegos para Steam

## Introducción

El presente informe detalla el proceso de desarrollo de un sistema de recomendación de videojuegos para la plataforma Steam, utilizando técnicas de aprendizaje automático. El proyecto se centra en el procesamiento de datos, análisis exploratorio y construcción de un modelo de recomendación basado en la similitud entre juegos.

## Procesamiento de Datos

### ETL (Extract, Transform, Load)

Se crearon tres programas de Python para realizar la limpieza y normalización de los datos de entrada, los cuales se encontraban en tres archivos JSON:

- `australian_user_reviews.json`
- `australian_users_items.json`
- `output_steam_games.json`

#### ETL\_user\_reviews

- Corrige errores de lectura y formato de datos.
- Normaliza tipos de datos y elimina columnas innecesarias.
- Aplica análisis de sentimiento a las reseñas de usuarios.
- Guarda los datos procesados en archivos CSV y Parquet.

#### ETL\_output\_steam\_games

- Elimina filas vacías y renombra columnas.
- Convierte valores de precio a tipo float.
- Transforma fechas de lanzamiento a formato datetime.
- Guarda los datos procesados en archivos CSV y Parquet.

#### ETL\_australian\_users\_items

- Corrige errores y carga datos desanidados.
- Convierte cadenas a listas y maneja valores NaN.
- Guarda los datos procesados en archivos CSV y Parquet.

## Análisis Exploratorio de Datos (EDA)

Se realizó un análisis exhaustivo de cada archivo CSV generado, incluyendo:

- Identificación y corrección de valores atípicos.
- Visualización de distribuciones y relaciones entre variables.
- Generación de estadísticas descriptivas y métricas relevantes.

### Ejemplos de EDA:

- Histogramas de variables numéricas.
- Gráficos de dispersión para relaciones entre variables.
- Análisis de frecuencia de etiquetas de juegos.

## Modelo de Machine Learning

Se implementó un modelo de recomendación basado en la similitud entre juegos utilizando el algoritmo de vecinos más cercanos (KNN). El proceso incluyó:

- Preprocesamiento de datos y extracción de características.
- Entrenamiento del modelo utilizando la similitud coseno.
- Generación de recomendaciones personalizadas para los usuarios.

### Ejemplo de Uso del Modelo:

- Se proporciona una función para obtener recomendaciones para un juego específico.

## Conclusiones

El proyecto ha logrado desarrollar un sistema de recomendación de juegos para la plataforma Steam, que proporciona recomendaciones personalizadas basadas en el análisis de datos y características de los juegos. El enfoque de procesamiento de datos, análisis exploratorio y modelado ha demostrado ser efectivo para crear un sistema útil y relevante para los usuarios de Steam.