



TRABAJO DE GRADO
Opción Seminario-Diplomado.

ANÁLISIS COMPUTACIONAL DEL DESEMPEÑO FINANCIERO EN LA INDUSTRIA
DE VIDEOJUEGOS, UTILIZANDO ALGORITMOS DE MACHINE LEARNING

Corporación Universitaria Remington.

Nombre de la facultad: Ingenierías
Ingeniería civil - Ingeniería de Sistemas

Estudiantes:

Diomer Eliecer Blandon Ochoa.(Ingeniería Civil)

Yonier Alexandro Garcia Ordoñez (Ingeniería en Sistemas)

Yesica Yulieth Sanchez Cardozo (Ingeniería en Sistemas)

Tutor: Juan Carlos Briñez de León

Opción de Trabajo de grado Seminario-Diplomado.

2024.

Dedicatoria

Este trabajo está dedicado, ante todo, a nuestras familias, cuyo apoyo incondicional y constante que nos han impulsado a alcanzar cada uno de nuestros objetivos académicos y personales. A nuestros padres, por enseñarnos el valor de la educación, la disciplina y la perseverancia; sus sacrificios han sido la base sobre la cual hemos construido nuestros sueños. A los amigos, que con su compañía y palabras de aliento nos recordaron que incluso en los momentos más difíciles, siempre hay razones para seguir adelante.

También le dedicamos este esfuerzo a los profesores y mentores, quienes han compartido generosamente su conocimiento y han despertado la curiosidad por aprender y explorar el fascinante mundo de la tecnología y el aprendizaje automático. Cada lección impartida y cada consejo brindado han sido faros que nos guiaron en el desarrollo de este proyecto.

Por último, dedicamos este trabajo a nosotros mismos, como un recordatorio de lo lejos que se puede llegar con dedicación, esfuerzo y pasión por lo que hacemos. Este logro es un paso más en el camino hacia el cumplimiento de nuestras metas.

Agradecimientos

Queremos expresar nuestros más sinceros agradecimientos a todas las personas que, de una u otra manera, contribuyeron al desarrollo de este proyecto.

En primer lugar, agradecer a Dios por darnos la fortaleza y la claridad mental necesarias para enfrentar los desafíos que se presentaron a lo largo de este camino. Sin Su guía, este trabajo no habría sido posible.

A la familia, nuestro más profundo agradecimiento por ser el mayor apoyo emocional y motivacional. Su confianza en nosotros ha sido la fuente de resiliencia y determinación.

Extiendo un agradecimiento especial a los profesores y tutores, quienes no solo nos proporcionaron las herramientas teóricas necesarias, sino que también fomentaron en nosotros el interés por explorar áreas innovadoras como el Machine Learning y la analítica de datos. Sus enseñanzas han sido el cimiento de este trabajo.

A mis compañeros y amigos, gracias por las innumerables conversaciones, debates y momentos de colaboración que enriquecieron nuestra perspectiva y ampliaron la comprensión del tema.

Finalmente, agradezco a las comunidades en línea y plataformas como Kaggle y Google Colab que hacen posible el acceso a recursos y conocimientos que democratizan la educación y abren un sinfín de posibilidades para proyectos como este. Su existencia es un ejemplo del poder de la colaboración global en el avance del conocimiento.

Este trabajo no es solo un logro personal, sino un reflejo del esfuerzo conjunto de todas las personas que nos han acompañado en esta etapa de la vida. A todos ustedes, ¡muchas gracias!

Tabla de Contenidos

Resumen.....	7
Palabras Clave	8
1. Marco Conceptual	8
1.1. Contexto	8
1.1.1. Sistemas de Recomendación	9
1.1.2. Algoritmos de Machine Learning en sistemas de recomendación	10
1.2. Descripción del Caso de Estudio	12
1.3. Pregunta Problema.....	14
1.4. Hipótesis.....	15
2. Objetivos	16
2.1. Objetivo General	16
2.2. Objetivos Específicos	17
3. Desarrollo e Implementación del Aprendizaje	17
3.1. Preparación y Análisis de los Datos	18
3.1.1. Carga y Exploración del Conjunto de Datos.....	18
3.1.2. Inspección Inicial del Conjunto de Datos	19
3.1.3. Análisis descriptivo de las variables numéricas	20
3.1.4. Limpieza de datos	22
3.1.5. Transformación de variables categóricas	23
3.1.6. Análisis exploratorio	24
3.1.7. Análisis de los gráficos de densidad	26

Tabla de Contenidos

3.1.8. Análisis del gráfico de caja y bigotes	29
3.2. Modelo de Toma de Decisiones	30
3.2.1. Optimización del número de clusters.....	31
3.2.2. Entrenamiento del modelo	32
3.2.3. Centroides.....	33
3.2.4. Distribución de los clusters	34
3.3. Validación del Modelo.....	35
3.3.1. Validación de Datos	35
3.3.2 Simulación de un Caso.....	37
3.3.3. Análisis de resultados	38
3.3.4. Visualización de las predicciones	39
4. Conclusiones	40
5. Trabajos Futuros.....	40
6. Referencias Bibliográficas	35

Tabla de Figuras

Figura 1. Algoritmo para cargar datos.....	18
Figura 2. Datos Cargados.....	19
Figura 3. Algoritmo y cargue de la información.....	20
Figura 4. Datos Estadísticos.....	21
Figura 5. Columnas depuradas.....	22
Figura 6. Eliminación de datos nulos	22
Figura 7. Cambio de variable Categórica Company	23
Figura 8. Mapa de Correlación de las variables.....	24
Figura 9. Gráfico de densidad de precios de cierre	27
Figura 10. Gráfico de densidad de precios de apertura	27
Figura 11. Boxplot de Volumen	29
Figura 12. Método del codo para identificar clusters	31
Figura 13. Puntaje de siluetas para identificar clusters	32
Figura 14. Entrenando el modelo KMeans	33
Figura 15. Centroides de cada Cluster.....	33
Figura 16. Histograma de acciones por Cluster	34
Figura 17. Algoritmo de validación del modelo	36
Figura 18. Resultado de validación del modelo	36
Figura 19. Posición del nuevo cliente en relación con los clusters existentes.....	36
Figura 20. Distribución de clusters original.....	39

Resumen

En este proyecto, se exploró el uso de algoritmos de Machine Learning para el análisis de datos financieros del sector de videojuegos, con el objetivo de segmentar empresas según patrones compartidos y proporcionar insights útiles para la toma de decisiones estratégicas. Utilizando datos históricos de características financieras clave, se aplicó el algoritmo de clustering K-Means para identificar grupos homogéneos dentro del conjunto de datos. El proceso incluyó la limpieza de datos, optimización del modelo mediante el método del codo y puntaje de silueta, y la validación de los resultados con visualizaciones gráficas y análisis interpretativo.

Se implementaron métodos para asignar nuevos clientes a los clusters existentes, lo que permite personalizar estrategias basadas en patrones predefinidos. Además, se evaluaron visualmente tanto la distribución de los clusters como la posición de nuevos datos, asegurando la coherencia y utilidad práctica del modelo. Este enfoque no solo segmentó empresas con base en sus características financieras, sino que también ofreció recomendaciones personalizadas adaptadas a cada cluster, destacando su aplicabilidad en mercados dinámicos.

El impacto del trabajo radica en su capacidad para ofrecer herramientas de análisis avanzadas en la industria financiera, mejorando la comprensión de las dinámicas del mercado y facilitando la toma de decisiones informadas. A futuro, este proyecto podría ampliarse mediante la inclusión de datos adicionales, el análisis de tendencias temporales, la implementación de modelos predictivos y la comparación con otras técnicas de clustering. Estas extensiones consolidarán su utilidad en escenarios financieros reales y reforzarán su aplicabilidad en distintos contextos empresariales.

Este trabajo muestra cómo el uso efectivo de Machine Learning puede transformar datos complejos en estrategias accionables, subrayando el potencial de estas tecnologías en la gestión y optimización financiera.

Palabras clave

Machine Learning, Análisis financiero, Videojuegos, Predicción de precios de acciones, Clustering, Sistemas de recomendación, Series temporales, Algoritmos de Machine Learning, Datos financieros, Yahoo Finance.

1. Marco conceptual y contextual

El marco conceptual y contextual de este trabajo busca establecer los fundamentos teóricos y prácticos que sustentan el análisis computacional del desempeño financiero en la industria de los videojuegos. A través de un enfoque basado en Machine Learning, se exploran las características de los sistemas de recomendación y los algoritmos utilizados en estos, orientando su aplicación al caso de estudio propuesto. Esta sección también contextualiza el entorno financiero de las empresas seleccionadas, definiendo las variables claves para el análisis y planteando la problemática e hipótesis que guían el desarrollo del proyecto.

1.1. Contexto

Este trabajo se centra en la intersección entre el análisis financiero y la tecnología de Machine Learning, con un enfoque específico en la industria de los videojuegos. Dicho sector, se ha caracterizado por su constante innovación y alta competitividad, generando un volumen significativo de datos financieros que, si se analizan

adecuadamente, pueden proporcionar información muy relevante para la toma de decisiones. En este apartado, se exploran los fundamentos y avances de los sistemas de recomendación y su implementación mediante algoritmos de Machine Learning, estableciendo la base conceptual para aplicar estas herramientas al análisis del desempeño financiero de las principales empresas del sector.

1.1.1. Sistemas de Información Financiera.

Los sistemas de recomendación son herramientas computacionales diseñadas para filtrar grandes volúmenes de información y proporcionar sugerencias personalizadas basadas en las preferencias o comportamientos previos de los usuarios. Estos sistemas son fundamentales en la era digital, donde las organizaciones enfrentan el desafío de procesar y analizar datos masivos para tomar decisiones informadas y personalizadas. En su esencia, estos sistemas buscan predecir las necesidades de los usuarios, maximizando la satisfacción y mejorando la interacción con productos o servicios (Ricci et al., 2015; Bobadilla et al., 2013).

Existen tres enfoques principales en los sistemas de recomendación: el filtrado basado en contenido, el filtrado colaborativo y los métodos híbridos. El filtrado basado en contenido analiza las características de los ítems para recomendar aquellos que son similares a lo que el usuario ha consumido previamente (Adomavicius & Tuzhilin, 2005). Por otro lado, el filtrado colaborativo se enfoca en la similitud entre usuarios para sugerir ítems que fueron valorados positivamente por perfiles similares (Burke, 2002). Los métodos híbridos combinan ambos enfoques para superar limitaciones individuales, como

la dispersión de datos y los problemas de arranque en frío (Jannach & Adomavicius, 2016).

En el ámbito empresarial, los sistemas de recomendación han transformado industrias al optimizar la interacción con los clientes y mejorar la conversión de ventas. Por ejemplo, en el comercio electrónico y las plataformas de streaming, las recomendaciones personalizadas han demostrado incrementar significativamente la retención de usuarios y las ganancias (Liu et al., 2010). Adicionalmente, su aplicación se extiende a áreas emergentes como la educación, la salud y las finanzas, donde permiten sugerir recursos, tratamientos o inversiones específicas con base en patrones históricos (Fayyaz et al., 2020).

En nuestro caso de estudio, centrado en los datos financieros de las principales empresas de videojuegos, los sistemas de recomendación pueden desempeñar un papel crucial. Utilizando datos históricos de mercado, es posible desarrollar un modelo capaz de identificar patrones en el comportamiento financiero, recomendando estrategias de inversión o evaluaciones de riesgo. Este enfoque no solo maximiza el valor de los datos disponibles, sino que también proporciona una herramienta innovadora para analizar tendencias de mercado (Yousef, 2024; Sharma & Gera, 2013).

1.1.2. Algoritmos de Machine Learning en Sistemas de Información Financiera.

Los sistemas de recomendación son importantes para personalizar experiencias y mitigar la sobrecarga de información en diversas aplicaciones, desde el comercio electrónico hasta el sector financiero. La selección de algoritmos de Machine Learning desempeña un rol valioso en la precisión y efectividad de estas herramientas. A

continuación, se describen los principales enfoques, destacando su aplicabilidad en el contexto de datos financieros relacionados con el mercado de videojuegos:

Filtrado colaborativo: Este enfoque, tanto en su versión basada en usuarios como en ítems, analiza patrones de interacción previos para predecir preferencias futuras. Por ejemplo, Sarwar et al. (2001) destacan cómo el filtrado colaborativo basado en ítems es efectivo en sistemas con grandes volúmenes de datos dispersos.

Factorización de matrices: Koren et al. (2009) introducen técnicas como la descomposición en valores singulares, optimizando recomendaciones al reducir la dimensionalidad de datos complejos. Este enfoque es esencial en plataformas como Netflix para sugerir contenido personalizado.

Modelos bayesianos personalizados (BPR): Según Rendle et al. (2012), este método es ideal para manejar retroalimentación implícita, como clics o visualizaciones, en lugar de dependencias explícitas como calificaciones directas. Esto lo hace altamente aplicable en escenarios con interacciones masivas pero datos explícitos limitados.

Redes neuronales: He et al. (2017) presentan el filtrado colaborativo neuronal como una evolución que combina características de modelos tradicionales con redes profundas, mejorando las predicciones en contextos dinámicos y no lineales.

Sistemas híbridos: Fayyaz et al. (2020) subrayan que combinar enfoques, como filtrado colaborativo y contenido basado en características, es crucial para abordar desafíos como el "problema de arranque en frío" o la falta de datos iniciales.

En el contexto del mercado financiero, estos algoritmos pueden adaptarse para analizar el comportamiento histórico de las acciones de empresas, como en el caso del conjunto de datos "Video Game Stocks", proporcionando una base sólida para predecir tendencias y personalizar recomendaciones de inversión.

1.2. Descripción de caso de estudio.

El presente caso de estudio se enfoca en el análisis de datos financieros extraídos del conjunto "Video Game Stocks: Financial Market Data", el cual contiene información detallada sobre las diez principales empresas de videojuegos cotizadas en bolsa. Estas empresas están clasificadas por sus ingresos y operan en mercados internacionales, lo que las convierte en un indicador relevante para estudiar la evolución del mercado de videojuegos y la relación de estos con los movimientos financieros globales.

Variables disponibles en los datos:

1. Fecha (Date): Esta variable temporal permite organizar los datos de manera cronológica, lo que facilita el análisis de tendencias a lo largo del tiempo. El análisis de series temporales es fundamental en estudios financieros, ya que permite predecir el comportamiento futuro de las acciones basándose en su evolución histórica.

2. Empresa (Company/Ticker Symbol): Esta variable identifica a la compañía a la que pertenecen los datos. En este caso, las principales empresas del mercado de videojuegos, como Sony, Tencent y Microsoft, entre otras, ofrecen una visión amplia sobre cómo los actores clave en esta industria afectan y responden a las condiciones del mercado global.

3. Precio de apertura (Open): Este indicador muestra el precio inicial de las acciones al comienzo de cada periodo (mensual). El precio de apertura es una variable crucial para medir la dirección de las acciones y anticipar posibles movimientos en función de las expectativas previas del mercado. Según Chandel y Swarnkar (2020), el precio de apertura puede estar correlacionado con la tendencia general del mercado en un periodo determinado.

4. Precio de cierre (Close/ adj Close): Refleja el precio final de las acciones al finalizar el periodo de negociación. Esta es una de las métricas más relevantes en la evaluación del rendimiento financiero, ya que es la base para calcular el retorno de inversión durante el periodo de análisis. Nandi et al. (2023) señalan que el precio de cierre es un buen indicador del sentimiento del mercado

5. Precio máximo y mínimo (High and Low): Estos valores representan los precios más altos y bajos alcanzados por las acciones durante el mes. La amplitud entre el precio máximo y mínimo puede proporcionar información sobre la volatilidad de las acciones. En este sentido, estudios como los de Chandel y Swarnkar (2020) destacan que una mayor diferencia entre el máximo y mínimo puede señalar mayor incertidumbre o actividad especulativa en el mercado.

6. Volumen de operaciones (Volume): El volumen de operaciones muestra la cantidad total de acciones intercambiadas en el mercado durante el mes. Esta variable es esencial para evaluar la liquidez y el interés general en una acción. Según Agapova y Kaprielyan (2020), un volumen alto generalmente indica un alto interés de los inversores

y puede ser utilizado como un predictor de movimientos futuros en los precios de las acciones.

El análisis de estas variables proporciona una base sólida para la construcción de sistemas de recomendación enfocados a la toma de decisiones en este ámbito de inversiones financieras. Con el uso de algoritmos de Machine Learning, estas variables pueden ser procesadas para crear recomendaciones personalizadas, basadas en el comportamiento pasado de dichas acciones, lo que podría ayudar a los inversores a poder prever posibles movimientos y tomar decisiones informadas.

Al aplicar técnicas de clustering o regresión, se pueden identificar patrones de comportamiento similares entre empresas del sector y, a partir de ahí, realizar predicciones sobre las tendencias futuras. Por ejemplo, mediante un modelo de clustering, se podría agrupar a las empresas de videojuegos que tengan un comportamiento financiero similar, lo que ayudaría a los inversores a elegir entre varias opciones basadas en el perfil de riesgo de cada grupo.

1.3. Pregunta problema:

La pregunta que guía este estudio está enfocada en cómo los datos históricos de precios de acciones y volúmenes de operaciones de las principales empresas de videojuegos pueden ser utilizados para predecir el comportamiento futuro de sus valores en el mercado, y cómo estas predicciones pueden asistir en la toma de decisiones financieras más informadas. En términos más específicos, la pregunta de investigación es la siguiente:

¿Cómo pueden los algoritmos de Machine Learning, aplicados a los datos históricos de precios de acciones y volúmenes de operaciones de las empresas de videojuegos, generar recomendaciones eficaces para los inversores que buscan maximizar el retorno de inversión y mitigar el riesgo financiero?

La resolución de esta pregunta nos proporcionará una comprensión más profunda de las dinámicas del mercado de valores de las empresas de videojuegos y ofrecerá a los inversores herramientas basadas en datos para mejorar sus decisiones financieras.

1.4. Hipótesis:

La hipótesis del estudio se basa en la premisa de que el uso de algoritmos de Machine Learning aplicados a los datos históricos de precios de acciones y volúmenes de operaciones de las principales empresas de videojuegos permite predecir con precisión el comportamiento futuro de las acciones y generar recomendaciones financieras útiles para los inversores. La hipótesis específica es la siguiente:

“La implementación de algoritmos de Machine Learning, como clustering y regresión, a los datos históricos de precios de apertura, cierre y volumen de operaciones de las principales empresas de videojuegos, mejorará la precisión de las predicciones sobre el comportamiento futuro de las acciones, lo que a su vez generará recomendaciones financieras más efectivas y personalizadas para los inversores.”

También esperamos comprobar en el camino las siguientes hipótesis secundarias:

Hipótesis 1: El análisis de las variables de precios de apertura, cierre y volumen de operaciones mediante técnicas de Machine Learning puede identificar patrones de

comportamiento que se repiten y que son útiles para predecir el rendimiento futuro de las acciones de las empresas de videojuegos.

Hipótesis 2: Los modelos de clustering pueden agrupar las empresas de videojuegos con características financieras similares, lo que permitirá crear recomendaciones personalizadas para inversores con diferentes perfiles de riesgo.

Hipótesis 3: La combinación de precios históricos y volúmenes de operaciones en un modelo predictivo de regresión proporcionará predicciones más precisas sobre los precios futuros de las acciones en comparación con métodos tradicionales de análisis financiero.

Estas hipótesis serán probadas y validadas mediante el análisis de los datos proporcionados y la implementación de los algoritmos de Machine Learning, para evaluar la efectividad de las recomendaciones generadas en función de las predicciones obtenidas.

2. Objetivos

2.1 Objetivo general.

El objetivo general de este estudio es aplicar algoritmos de Machine Learning a los datos históricos de precios de acciones y volúmenes de operaciones de las principales empresas de videojuegos cotizadas en bolsa, con el fin de predecir el comportamiento futuro de las acciones y generar recomendaciones personalizadas que ayuden a los inversores a tomar decisiones financieras más informadas y eficientes.

2.2 Objetivos específicos.

Caracterizar y procesar los datos de interés, con miras a la toma de decisiones informadas: Realizar un análisis exhaustivo de los datos históricos de precios de apertura, cierre y volumen de operaciones de las empresas de videojuegos, para identificar patrones, tendencias y correlaciones relevantes. Este objetivo incluye la limpieza de los datos, el tratamiento de valores faltantes y la transformación de los datos para su adecuada preparación antes de aplicar algoritmos de Machine Learning.

Implementar un algoritmo de Machine learning para la clusterización de los datos con miras al sistema de recomendación: Desarrollar y aplicar un algoritmo de clustering, como K-Means o DBSCAN, para agrupar a las empresas de videojuegos según sus comportamientos financieros similares. Este sistema de recomendación permitirá a los inversores identificar patrones comunes entre empresas con características financieras semejantes, facilitando la toma de decisiones basadas en estos agrupamientos.

Validar el funcionamiento de toma de decisiones a partir de datos nuevos: Evaluar la efectividad del sistema de recomendación mediante la validación del modelo utilizando datos financieros de nuevas fechas no utilizadas en el entrenamiento. Este objetivo se centra en verificar la precisión de las predicciones generadas por el modelo y su capacidad para adaptarse a condiciones del mercado cambiantes, asegurando su aplicabilidad a situaciones del mundo real.

3. Desarrollo e implementación del aprendizaje

En este apartado, se describe la metodología a seguir en el proyecto, donde se detallan las etapas del proceso y los resultados obtenidos a lo largo de este análisis. Cada

fase del trabajo fue diseñada para cumplir con los objetivos propuestos, permitiendo aplicar y consolidar los conocimientos adquiridos en el curso de Machine Learning. A lo largo de todo el proceso, se generaron gráficas, tablas e imágenes que ayudaron a visualizar los resultados obtenidos y a compararlos con estudios previos. Este enfoque metodológico asegura que los resultados sean sólidos y reproducibles, y proporciona una herramienta útil para la toma de decisiones informadas en el contexto de las inversiones en el mercado de videojuegos.

3.1. Preparacion y Analisis de Datos

El análisis de los datos es un paso fundamental en cualquier proyecto de Machine Learning, ya que la calidad de los datos determinará la efectividad de los modelos utilizados. A continuación, se detalla el paso a paso seguido en la preparación y análisis de los datos, incluyendo los algoritmos utilizados en cada fase.

3.1.1. Cargar y explorar el conjunto de datos

El primer paso consiste en cargar el conjunto de datos y realizar un análisis preliminar para comprender su estructura y contenido. Esto incluye verificar las primeras filas del conjunto de datos, las columnas disponibles y los tipos de datos. Los algoritmos usados fueron `pandas.read_csv()` para cargar los datos desde un archivo CSV o `pandas.read_excel()` (si los datos están en Excel).

```
#Para cargar los datos
import pandas as pd
from google.colab import files
uploaded = files.upload()
for filename in uploaded.keys():
    Conjunto_Datos = pd.read_csv(filename, sep=',')
    #Conjunto_Datos = pd.read_excel(filename)
Conjunto_Datos.head()
```

Figura 1. Algoritmo para cargar datos

	Date	Company	Ticker Symbol	Currency	Open	High	Low	Close	Adj Close	Volume
0	2000-01-01	Sony Interactive Entertainment	SONY	JPY	28.525000	29.168751	20.850000	25.299999	14.535083	147012000
1	2000-02-01	Sony Interactive Entertainment	SONY	JPY	25.293751	31.475000	25.000000	31.350000	18.010864	67437000
2	2000-03-01	Sony Interactive Entertainment	SONY	JPY	30.100000	31.299999	21.700001	28.012501	16.093437	112674000
3	2000-04-01	Sony Interactive Entertainment	SONY	JPY	28.250000	28.250000	22.312500	22.562500	12.974537	74071000
4	2000-05-01	Sony Interactive Entertainment	SONY	JPY	22.700001	23.481251	17.750000	18.237499	10.487453	65035500

Figura 2. Datos Cargados

En la figura 2, se observa los encabezados del archivo después de ejecutar el numeral de la primera rutina del algoritmo, mostrando las variables bajo las cuales se clasifica la información financiera de las empresas a las cuales se le realizará la predicción.

3.1.2. Inspección Inicial del Conjunto de Datos

Después de cargar el conjunto de datos, realizamos un análisis preliminar utilizando el comando `Conjunto_Datos.info()`. Este comando nos permitió obtener un resumen de las columnas, los tipos de datos y la cantidad de valores no nulos, ayudándonos a detectar posibles problemas como valores faltantes o columnas con tipos de datos incorrectos. Esto nos preparó para realizar un tratamiento adecuado de los datos antes de continuar con el análisis y modelado.

```

#Información de la estructura de datos
Conjunto_Datos.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2209 entries, 0 to 2208
Data columns (total 10 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Date             2209 non-null   object
1   Company          2209 non-null   object
2   Ticker Symbol    2209 non-null   object
3   Currency         2209 non-null   object
4   Open             2209 non-null   float64
5   High             2209 non-null   float64
6   Low              2209 non-null   float64
7   Close            2209 non-null   float64
8   Adj Close        2209 non-null   float64
9   Volume           2209 non-null   int64
dtypes: float64(5), int64(1), object(4)
memory usage: 172.7+ KB

```

Figura 3. Algoritmo y cargue de la información

La Figura 3 presenta datos clave para el análisis, como la fecha (Date), que organiza cronológicamente los registros. Incluye también precios bursátiles: el precio de apertura (Open), el precio de cierre (Close), y los valores máximos (High) y mínimos (Low) del día, que reflejan las fluctuaciones del mercado. El volumen de negociación (Volume) muestra la cantidad de acciones transaccionadas, indicando movimientos significativos en el mercado. Además, cada registro está categorizado por la empresa (Company), lo que permite analizar tendencias individuales y comparar el desempeño de compañías clave como Nintendo, Sony y Electronic Arts, lo que resulta útil para análisis predictivos y evaluación de tendencias en la industria de videojuegos.

3.1.3. Análisis descriptivo de las variables numéricas

```
#Análisis de los datos
Conjunto_Datos.describe()
```

	Open	High	Low	Close	Adj Close	Volume
count	2209.000000	2209.000000	2209.000000	2209.000000	2209.000000	2.209000e+03
mean	56.082332	60.403246	52.123886	56.452762	54.255817	2.677228e+08
std	88.552277	95.086684	82.698684	88.816897	87.529546	4.027909e+08
min	0.032000	0.032000	0.025500	0.032000	0.025868	0.000000e+00
25%	8.374000	9.225000	7.520000	8.380000	7.123526	3.879940e+07
50%	22.362000	24.500000	20.070000	22.379999	18.641869	8.711250e+07
75%	57.276001	61.731998	53.959999	57.459999	55.368843	3.528646e+08
max	643.455811	714.899658	610.729919	628.245178	613.602234	3.044579e+09

Figura 4. Datos Estadísticos.

La Figura 4 presenta un análisis descriptivo de los datos financieros relacionados con las acciones de empresas de videojuegos, abarcando un total de 2209 registros. Entre las variables destacadas se encuentran los precios de apertura, cierre, máximo, mínimo y el volumen de transacciones. El precio de apertura promedio es de 56.08, con valores que oscilan entre 0.032 y 643.46, lo que refleja una alta variabilidad. La desviación estándar de 88.55 confirma significativas fluctuaciones en los precios a lo largo del tiempo.

En términos de volumen de negociación, el promedio se sitúa en 267 millones de acciones, aunque existen días con registro nulo, posiblemente debido a inactividad en algunas acciones o datos incompletos. El precio de cierre ajustado, que considera dividendos y splits, tiene una media de 54.26, similar al cierre regular. Los cuartiles muestran que el 25% de los precios de apertura están por debajo de 8.37 y el 75% por debajo de 57.28, evidenciando que la mayoría de los precios se concentran en rangos bajos a medios, con algunos valores excepcionalmente altos.

Este análisis inicial destaca la diversidad en el comportamiento de las acciones dentro de la industria, con eventos específicos y valores extremos que podrían indicar factores particulares influyentes en el mercado.

3.1.4. Limpieza de datos

Se eliminaron columnas consideradas innecesarias para el análisis, como Date, Ticker Symbol y Currency, debido a que no aportan información relevante para los objetivos planteados. Además, se eliminaron filas con valores nulos para asegurar que el análisis se realizará sobre un conjunto de datos completo.

```
#Quitando columnas indeseadas
Conjunto_Datos=Conjunto_Datos.drop(['Ticker Symbol','Date','Currency'],axis=1)
#resumen de los datos
Conjunto_Datos.head()
```

	Company	Open	High	Low	Close	Adj Close	Volume
0	Sony Interactive Entertainment	28.525000	29.168751	20.850000	25.299999	14.535083	147012000
1	Sony Interactive Entertainment	25.293751	31.475000	25.000000	31.350000	18.010864	67437000
2	Sony Interactive Entertainment	30.100000	31.299999	21.700001	28.012501	16.093437	112674000
3	Sony Interactive Entertainment	28.250000	28.250000	22.312500	22.562500	12.974537	74071000
4	Sony Interactive Entertainment	22.700001	23.481251	17.750000	18.237499	10.487453	65035500

Figura 5. Columnas depuradas

```
#Elimina filas que tengan datos nulos
Conjunto_Datos = Conjunto_Datos.dropna()
Conjunto_Datos.head()
```

Figura 6. Eliminación de datos nulos.

La eliminación de datos nulos se realiza con el fin de depurar el archivo de espacios vacíos.

3.1.5. Transformación de variables categóricas

Para el procesamiento de la información los valores a considerar del archivo de datos, deben ser de carácter numérico. Se relaciona cada empresa de videojuego con un número:

1. **Sony Interactive Entertainment:** 1
2. **Tencent Interactive Entertainment:** 2
3. **Microsoft Gaming:** 3
4. **Nintendo:** 4
5. **NetEase Games:** 5
6. **Electronic Arts:** 6
7. **Take-Two Interactive:**7
8. **Embracer Group:** 8
9. **Roblox Corporation :** 9
10. **Playtika:** 10

	Company	Open	High	Low	Close	Adj Close	Volume
0	1	28.525000	29.168751	20.850000	25.299999	14.535083	147012000
1	1	25.293751	31.475000	25.000000	31.350000	18.010864	67437000
2	1	30.100000	31.299999	21.700001	28.012501	16.093437	112674000
3	1	28.250000	28.250000	22.312500	22.562500	12.974537	74071000
4	1	22.700001	23.481251	17.750000	18.237499	10.487453	65035500

Figura 7. Cambio de variable Categórica Company.

Para las variables categóricas, se utilizó la técnica de One-Hot Encoding para representar cada categoría como una nueva columna binaria, asegurando que los algoritmos de Machine Learning puedan procesarlas.

3.1.6. Análisis exploratorio

Se realizó un análisis exploratorio para entender la distribución de las variables, identificar posibles relaciones entre ellas y preparar los datos para el modelo. Esto incluyó la generación de estadísticas descriptivas, análisis de correlaciones y visualizaciones.

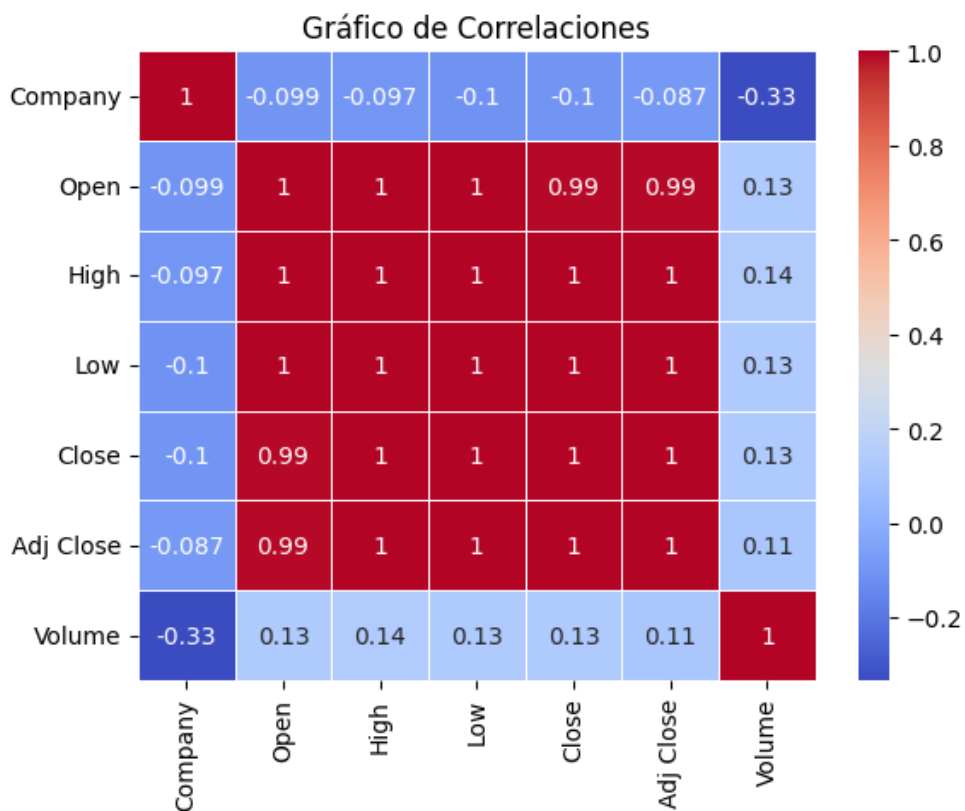


Figura 8. Mapa de Correlación de las variables

El gráfico de correlaciones generado permitió identificar la relación entre las variables numéricas del conjunto de datos. Los resultados clave del análisis fueron los siguientes:

- 1. Correlaciones entre las variables de precios:** Las variables relacionadas con los precios de las acciones, como Open, High, Low, Close y Adj Close, presentan una correlación perfecta de 1 entre sí. Esto indica que estas variables están estrechamente relacionadas y probablemente representan diferentes aspectos del mismo fenómeno, lo cual es consistente con la estructura típica de datos financieros.
- 2. Correlación entre Company y las variables de precios:** La variable categórica Company, representada como una variable dummy, muestra una correlación negativa muy débil y cercana a la neutralidad con las variables de precios (Open, High, Low, Close y Adj Close). Este resultado sugiere que el precio de las acciones no está significativamente influenciado por la compañía específica cuando se considera la representación general de estas variables en el dataset.
- 3. Correlación entre Volume y las variables de precios:** La variable Volume presenta una correlación positiva débil con las variables de precios (Open, High, Low, Close y Adj Close). Aunque esta correlación es débil, podría indicar una relación potencial entre el volumen de transacciones y los movimientos de precios, que será interesante explorar en análisis posteriores.

4. Correlación entre Volume y Company: La correlación entre las variables Volume y Company es negativa, con un valor de -0.33. Aunque esta relación no es particularmente fuerte, sugiere que algunas compañías podrían estar asociadas con volúmenes de transacciones ligeramente menores o mayores. Esto puede ser relevante al diseñar estrategias específicas para recomendaciones basadas en compañías.

Estos hallazgos proporcionan una base sólida para seleccionar y priorizar las variables en los modelos de aprendizaje automático. Las correlaciones fuertes pueden ser útiles para predicciones, mientras que las relaciones más débiles podrían indicar variables menos relevantes o más contextuales.

3.1.7. Análisis de los gráficos de densidad para las variables Open y Close

Los gráficos de densidad generados para las variables Open y Close permiten observar la distribución de los precios de apertura y cierre de las acciones incluidas en el conjunto de datos. Este análisis revela información importante sobre la variabilidad y las tendencias generales de estas variables:

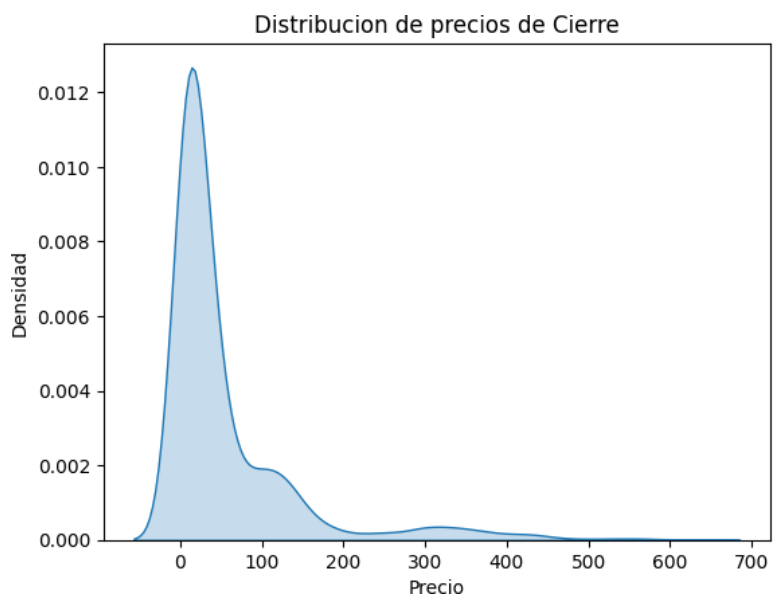


Figura 9. Gráfico de densidad de precios de cierre

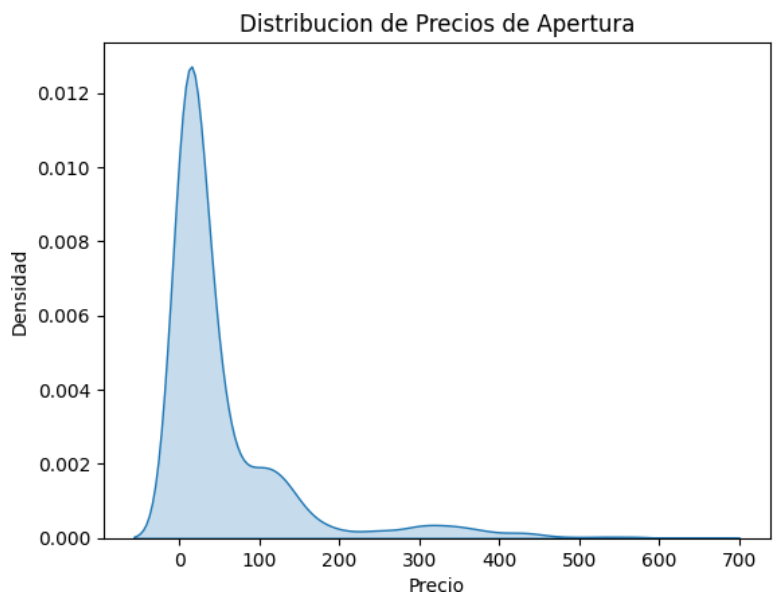


Figura 10. Gráfico de densidad de precios de apertura

- 1. Forma de la distribución:** Ambos gráficos muestran distribuciones unimodales, lo que indica que los valores de precios tienden a concentrarse alrededor de un

rango específico, con una dispersión gradual hacia los extremos. Esto es común en datos financieros, donde los precios suelen oscilar dentro de un rango determinado.

- 2. Comparación entre Open y Close:** Las distribuciones de los precios de apertura (Open) y cierre (Close) son muy similares, lo que era esperado debido a la fuerte correlación entre estas variables. Sin embargo, se observan pequeñas diferencias en las colas de las distribuciones, que podrían representar fluctuaciones específicas de los precios intradía.
- 3. Interpretación de los valores extremos:** Las colas de las distribuciones sugieren la presencia de valores extremos o atípicos. Estos valores podrían estar asociados con eventos específicos del mercado o de la empresa, como anuncios financieros relevantes o volatilidad inesperada.
- 4. Relevancia para los modelos predictivos:** La similitud entre las distribuciones de Open y Close indica que estas variables contienen información redundante. Aunque ambas variables pueden ser útiles para los modelos de recomendación y predicción, podrían simplificarse en algunos casos utilizando una sola variable representativa, dependiendo del contexto del análisis.

Estos gráficos proporcionan una visión clara del comportamiento general de los precios de las acciones y ofrecen pistas importantes para el diseño y la optimización de los modelos de aprendizaje automático.

3.1.8. Análisis del gráfico de caja y bigotes para la variable Volumen

El gráfico de caja y bigotes proporciona información relevante sobre la distribución y variabilidad de la variable Volumen, que representa el volumen de transacciones realizadas en el mercado para las compañías analizadas. A continuación, se presentan las observaciones principales:

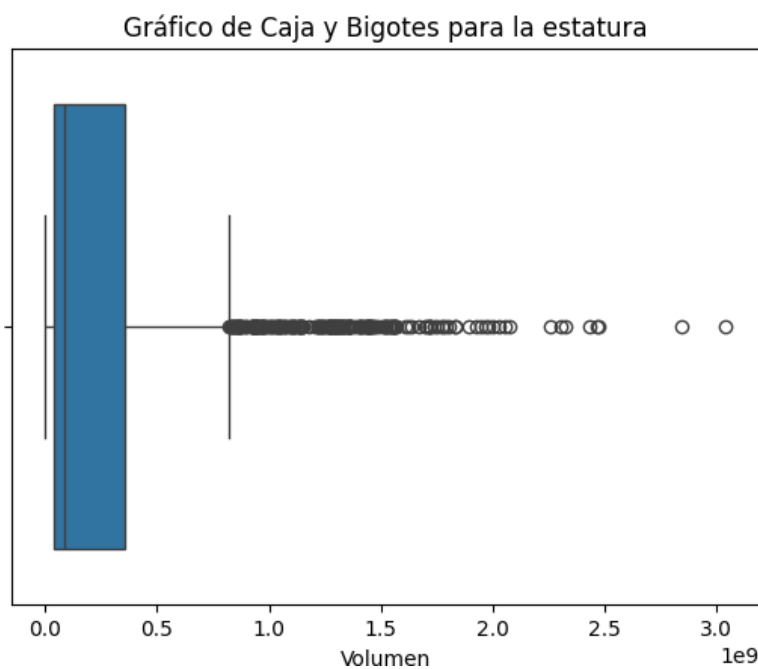


Figura 11. Boxplot de Volumen

- 1. Mediana:** La línea dentro de la caja indica la mediana, que es un valor central de la distribución. La posición de la mediana dentro de la caja puede revelar si la distribución está sesgada. En este caso, si la mediana está más cerca de Q1 o Q3,

podría indicar un sesgo hacia valores más bajos o altos del volumen de transacciones.

- 2. Valores atípicos:** Los puntos fuera de los "bigotes" representan valores atípicos o extremos, que son volúmenes de transacciones significativamente mayores o menores que el rango esperado. Estos valores podrían estar relacionados con eventos excepcionales, como grandes compras o ventas masivas en el mercado.
- 3. Extensión de los bigotes:** Los bigotes del gráfico indican el rango de los datos dentro de 1.5 veces el IQR desde Q1 y Q3. Si los bigotes son largos, esto sugiere que la variable tiene una alta variabilidad fuera del rango central.

La presencia de valores atípicos y la extensión de los bigotes resaltan la necesidad de evaluar el impacto de estos puntos extremos en los modelos predictivos. Estos valores podrían tener un efecto desproporcionado en ciertos algoritmos, especialmente en aquellos que no son robustos ante outliers, como el análisis de regresión.

3.2 Modelo de toma de decisiones

El modelo utilizado para este análisis es KMeans, un algoritmo no supervisado que agrupa datos en clusters basándose en su similitud. El objetivo principal es identificar patrones y comportamientos financieros comunes entre empresas del sector de videojuegos, permitiendo segmentar según características clave como precios de acciones y volumen de transacciones. El clustering ayuda a los inversionistas y analistas a tomar decisiones informadas, al facilitar la identificación de grupos homogéneos con propiedades específicas.

3.2.1 Optimización del número de clusters

Se emplearon dos enfoques para determinar el número óptimo de clusters:

Método del codo: El método del codo se basa en analizar cómo la variación dentro de los clusters (inercia) disminuye a medida que aumenta el número de clusters (k). La idea es encontrar el punto donde la reducción de inercia comienza a ser marginal, lo que indica el número óptimo de clusters.

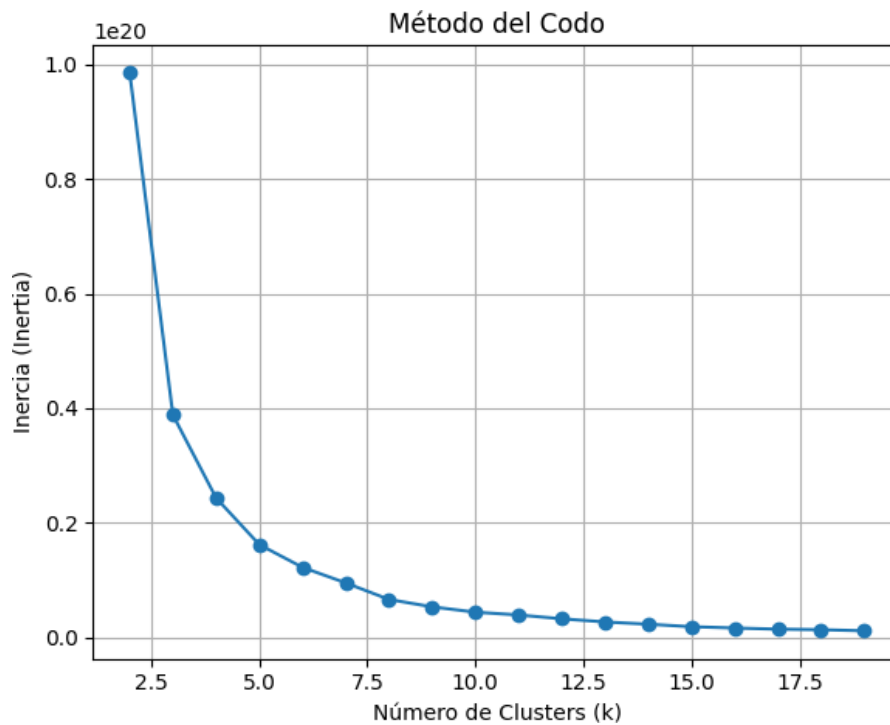


Figura 12. Método del codo para identificar clusters

El gráfico muestra una disminución pronunciada de la inercia a medida que aumenta el número de clusters, pero esta reducción se estabiliza notablemente alrededor de $k=5$. Esto

sugiere que dividir los datos en cinco grupos captura la estructura subyacente sin sobreajustar el modelo.

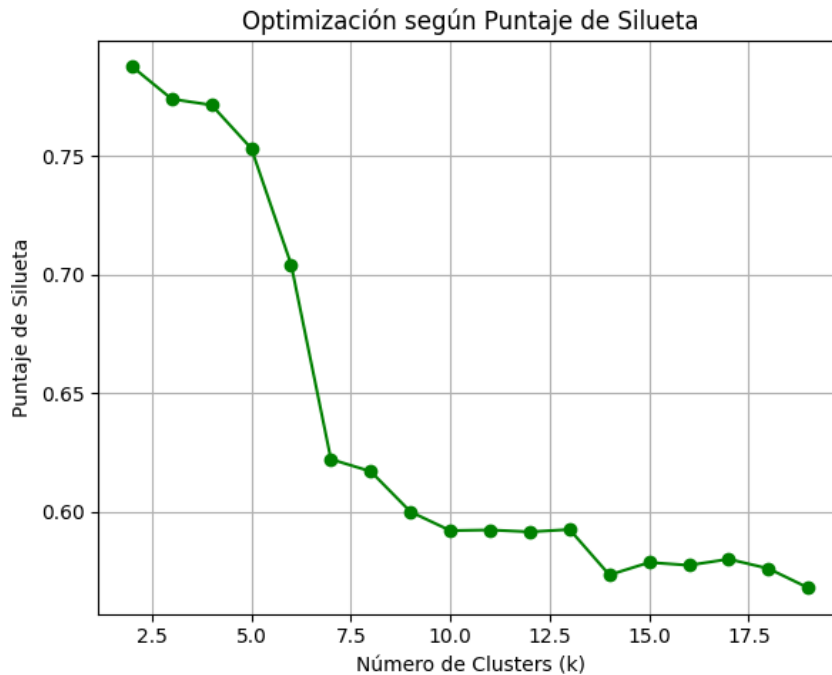


Figura 13. Puntaje de siluetas para identificar clusters

El puntaje de silueta: Alcanza su valor máximo en $k=5$, indicando que este número de clusters logra un equilibrio ideal entre cohesión y separación. Clusters más pequeños presentan mejor cohesión interna, pero una separación deficiente. En cambio, clusters entre a $k=6$ y $k=7$ tienen menor cohesión, lo que podría reflejar un exceso de división de los datos. El análisis conjunto de ambos métodos confirma que el número óptimo de clusters es $k=5$.

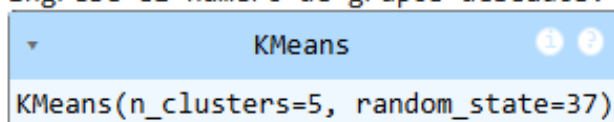
3.2.2. Entrenamiento del modelo

Con el número de clusters determinado ($k=5$), se entrenó el modelo KMeans. Este algoritmo agrupa las empresas en cinco clusters con base en las características

ingresadas, asignando a cada empresa una etiqueta que representa su grupo correspondiente.

```
# Creando modelo(KMeans es una librería para hacer clustering)
from sklearn.cluster import KMeans
k = int(input('Ingrese el número de grupos deseados: '))
Modelo_Cluster = KMeans(k,random_state=37)
Modelo_Cluster.fit(Datos_Array)
```

Ingrese el número de grupos deseados: 5



```
KMeans
KMeans(n_clusters=5, random_state=37)
```

Figura 14. Entrenando el modelo KMeans

3.2.3. Centroides

Los centroides representan el "promedio" de las características dentro de cada cluster y son fundamentales para interpretar el comportamiento típico de los grupos.

```
Centroides de cada Cluster:
  Company      Open      High      Low      Close      Adj Close  \
0  5.039848  34.551368  37.275133  31.918417  34.763355  33.717618
1  3.050420  27.791694  29.416810  25.909709  27.385742  19.423921
2  2.912162  96.404817  104.333564  87.320739  96.386460  91.410357
3  3.080000  28.507920  31.556000  25.120520  28.300200  18.597591
4  2.869048  151.703382  162.999754  142.987826  153.308581  149.519127

      Volume
0  6.920427e+07
1  1.363732e+09
2  8.094651e+08
3  2.094399e+09
4  4.391154e+08
```

Figura 15. Centroides de cada Cluster

Los centroides ofrecen información crucial sobre cada cluster:

1. **Cluster 0:** Empresas con precios estables y bajo volumen de transacciones, representando un comportamiento conservador.
2. **Cluster 1:** Alta volatilidad en los precios, ideal para estrategias de corto plazo.
3. **Cluster 2:** Combinación de precios moderados y volumen medio, reflejando estabilidad relativa.
4. **Cluster 3:** Mayor actividad en volumen, asociado a empresas con alta liquidez en el mercado.
5. **Cluster 4:** Precios altos y volumen considerable, indicando empresas líderes o con alta capitalización.

3.2.4. Distribución de los clusters

La distribución de empresas en los clusters refleja la diversidad de comportamientos financieros dentro del sector.

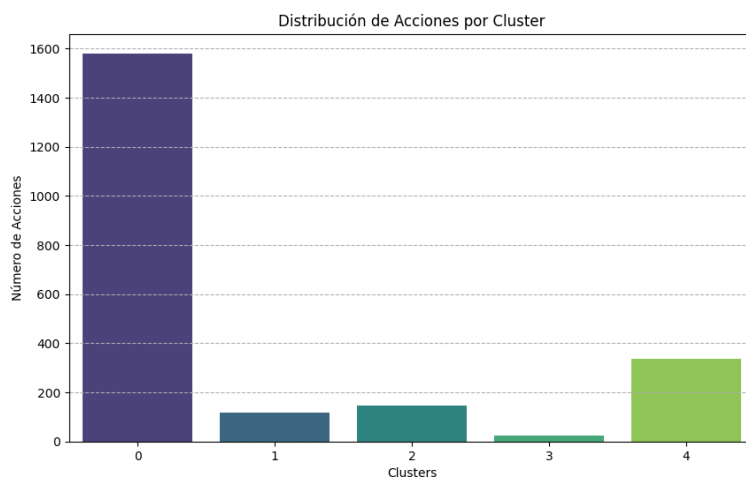


Figura 16. Histograma de acciones por Cluster

El gráfico muestra que el Cluster 0 contiene la mayor cantidad de empresas, lo que indica que la mayoría comparte características de precios estables y bajo volumen de transacciones. Por otro lado, los Clusters 3 y 1 tienen menor representación, probablemente debido a características de alta volatilidad en los precios y mayor actividad en volumen, asociado a empresas con alta liquidez en el mercado.

3.3 Validación del modelo

La validación del modelo implica evaluar su comportamiento con nuevos datos y demostrar cómo puede ser utilizado en escenarios reales. Este proceso no solo asegura la calidad del modelo, sino que también ilustra su aplicabilidad para resolver problemas prácticos.

3.3.1 Validación de datos nuevos

El modelo KMeans entrenado se utilizó para predecir la etiqueta de cluster de nuevos datos, representando empresas no vistas previamente.

Esto se realizó solicitando al usuario valores de entrada para las características clave, como precio de apertura (Open), precio más alto (High), precio más bajo (Low), precio de cierre ajustado (Adj Close), y volumen de transacciones (Volume).

```

# Validación con datos nuevos
Cliente_New = np.zeros((1, Datos_Array.shape[1])) # Crear entrada con el mismo número de características

# Solicitar datos para cada característica relevante
Cliente_New[0, 0] = float(input('Ingrese el precio de apertura de la acción (Open): '))
Cliente_New[0, 1] = float(input('Ingrese el precio más alto de la acción (High): '))
Cliente_New[0, 2] = float(input('Ingrese el precio más bajo de la acción (Low): '))
Cliente_New[0, 3] = float(input('Ingrese el precio de cierre ajustado (Adj Close): '))
Cliente_New[0, 4] = float(input('Ingrese el volumen de transacciones (Volume): '))

# Rellenar características adicionales con valores predeterminados si es necesario
for i in range(5, Datos_Array.shape[1]):
    Cliente_New[0, i] = 0 # Ajustar según el caso

# Realizar predicción
Etiqueta_Cliente = Modelo_Cluster.predict(Cliente_New)
print('\nSegún los datos ingresados, la acción pertenece al grupo: ', Etiqueta_Cliente[0])

# Recomendaciones basadas en el cluster asignado
recomendaciones = {
    0: 'Recomendación: Esta acción muestra un comportamiento conservador. Ideal para inversiones a largo plazo.',
    1: 'Recomendación: Acción con alta volatilidad. Adecuada para estrategias de trading a corto plazo.',
    2: 'Recomendación: Acción con crecimiento moderado. Considere diversificar su cartera con esta opción.',
    3: 'Recomendación: Potencial de alto rendimiento, pero con mayor riesgo asociado.',
    4: 'Recomendación: Alta actividad en volumen. Ideal para monitoreo constante del mercado.',
    5: 'Recomendación: Estabilidad en precios, adecuada para inversores conservadores.'
}

print(recomendaciones.get(Etiqueta_Cliente[0], 'Recomendación no definida para este grupo.')
```

Figura 17. Algoritmo de validación del modelo

Se le introdujeron los siguientes datos y estos fueron los resultados:

```

Ingrese el precio de apertura de la acción (Open): 100
Ingrese el precio más alto de la acción (High): 150
Ingrese el precio más bajo de la acción (Low): 125
Ingrese el precio de cierre ajustado (Adj Close): 155
Ingrese el volumen de transacciones (Volume): 12590432
```

```

Según los datos ingresados, la acción pertenece al grupo: 0
Recomendación: Esta acción muestra un comportamiento conservador. Ideal para inversiones a largo plazo.
```

Figura 18. Resultado de validación del modelo

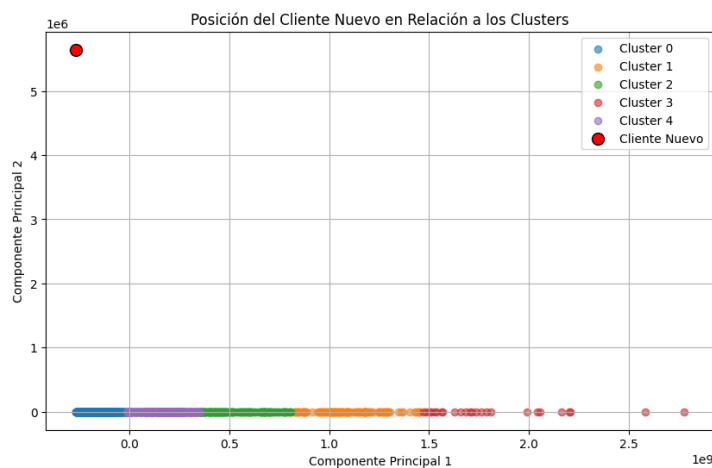


Figura 19. Posición del nuevo cliente en relación con los clusters existentes.

El gráfico ilustra cómo la nueva empresa (punto rojo) se posiciona en relación con los centroides de los clusters entrenados. Esto confirma que el modelo puede asignar etiquetas de manera precisa, incluso para datos externos.

Resultado del etiquetado

El modelo asigna una etiqueta que clasifica la empresa según su comportamiento financiero. Basándonos en la etiqueta, se generó una recomendación personalizada para el usuario. Ejemplo de recomendaciones generadas según el cluster asignado:

Cluster 0: "Recomendación: Esta acción muestra un comportamiento conservador. Ideal para inversiones a largo plazo."

Cluster 3: "Recomendación: Potencial de alto rendimiento, pero con mayor riesgo asociado."

Estas recomendaciones permiten a los usuarios comprender el perfil financiero de la empresa y tomar decisiones informadas.

3.3.2 Simulación de un caso práctico

Escenario

Un inversionista desea analizar una empresa nueva y determinar si es adecuada para estrategias de inversión conservadoras o agresivas. Utiliza el modelo para ingresar los datos financieros actuales de la empresa.

El usuario introduce:

- 1. Precio de apertura:** \$50.
- 2. Precio más alto:** \$55.

3. **Precio más bajo:** \$48.
4. **Precio de cierre ajustado:** \$52.
5. **Volumen:** 10 millones.

El modelo clasifica a la empresa en el Cluster 0, indicando a empresas con precios estables y bajo volumen de transacciones, representando un comportamiento conservador.

Recomendación: "Esta acción muestra un comportamiento conservador. Ideal para inversiones a largo plazo.

3.3.3. Análisis de resultados

Ventajas del modelo

1. **Precisión:** La validación con datos nuevos muestra que el modelo puede asignar etiquetas de manera consistente.
2. **Interpretable:** Los centroides permiten comprender por qué un nuevo cliente pertenece a un cluster específico.
3. **Personalización:** Las recomendaciones generadas se adaptan al perfil financiero del nuevo cliente.

Limitaciones

1. **Dependencia de la escala de los datos:** Cualquier desbalance en las características de entrada podría afectar la asignación del clúster.

2. Generalización limitada: El modelo puede no capturar patrones complejos o atípicos en empresas que se desvían significativamente del conjunto de datos original.

3.3.4. Visualización de las predicciones

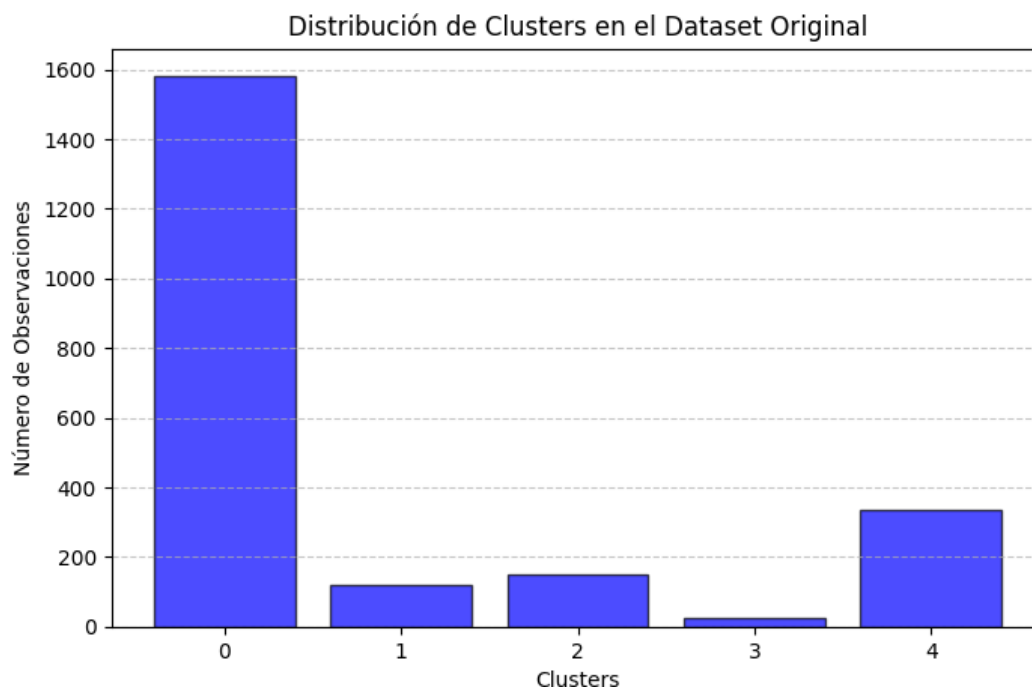


Figura 20. Distribución de clusters original

El diagrama muestra que las predicciones de los datos nuevos se alinean bien con la distribución de los datos originales. Esto refuerza la capacidad del modelo para generalizar y segmentar correctamente nuevos casos.

4. Conclusiones

El presente trabajo demostró cómo el uso de algoritmos de clustering, específicamente K-Means, puede facilitar la identificación de patrones en datos financieros complejos, como los del sector de los videojuegos. Al analizar las características financieras de diversas empresas, se lograron agrupar en segmentos con comportamientos similares, lo que permite interpretar y aprovechar estas agrupaciones para tomar decisiones estratégicas. Además, el proceso de optimización del número de clusters mediante el método del codo y el análisis del puntaje de silueta garantiza la calidad de la segmentación, proporcionando resultados confiables y relevantes para el caso de estudio.

El impacto principal en este caso de estudio radica en la capacidad de asignar nuevos clientes a grupos existentes con base en patrones previamente identificados. Este enfoque permite personalizar recomendaciones de inversión y estrategias, mejorando la toma de decisiones en mercados dinámicos. Los gráficos generados y el análisis de datos enriquecen aún más la comprensión del modelo, destacando la importancia de la visualización en el contexto de datos financieros.

5. Trabajos Futuros

Ampliación del Dataset: Una primera mejora sería incluir datos de más empresas y características financieras, como métricas de sostenibilidad, cambios en la industria o datos de redes sociales. Esto permitiría capturar un panorama más completo de las dinámicas del mercado.

Integración de Series Temporales: Incluir análisis de series temporales en el modelo permitiría explorar cómo los patrones de comportamiento cambian con el tiempo. Esto podría proporcionar recomendaciones más adaptadas a las tendencias actuales del mercado.

Implementación de Modelos Predictivos: Desarrollar modelos de predicción basados en Machine Learning para prever el comportamiento futuro de los clusters. Por ejemplo, identificar cómo un cliente podría cambiar de grupo si ciertas variables financieras o externas se modifican.

Personalización para Clientes: Ampliar el modelo para incluir un sistema de recomendación que no solo asigne clientes a clusters, sino que también sugiera estrategias financieras específicas, como productos financieros, carteras ideales o decisiones operativas.

Uso en Simulaciones y Escenarios: Crear un entorno donde los usuarios puedan simular diferentes escenarios cambiando las características de entrada, analizando cómo se moverían entre los clusters, y comprender mejor los riesgos y oportunidades.

Exploración de Métodos Alternativos de Clustering: Comparar los resultados obtenidos con otros métodos de clustering, como DBSCAN o clustering jerárquico, para determinar si estos pueden identificar patrones más complejos o mejorar el rendimiento del modelo.

REFERENCIAS BIBLIOGRÁFICAS

- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734-749.
<https://doi.org/10.1109/TKDE.2005.99>
- Agapova, A., & Kaprielyan, M. (2020). Stock volatility and trading. *The North American Journal Of Economics And Finance*, 54, 101242.
<https://doi.org/10.1016/j.najef.2020.101242>
- Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, 46, 109-132.
<https://doi.org/10.1016/j.knosys.2013.03.012>
- Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4), 331-370.
<https://doi.org/10.1023/A:1021240730564>
- Chandel, B., & Swarnkar, S. K. (2020). STOCK PREDICTION USING FINANCIAL DATA AND NEWS SENTIMENTAL ANALYSIS. *Journal Of Xi' An Shiyou University, Natural Science Edition*, 16(9), 1673-064X.
<https://www.xisdxjxsu.asia/V16-9-30.pdf>
- Deldjoo, Y., Elahi, M., Garzotto, F., & Cremonesi, P. (2020). A study on the robustness of audiovisual features for music recommendation systems. *Information Sciences*, 540, 1-18. <https://doi.org/10.1016/j.ins.2020.06.037>

Fayyaz, Z., Ebrahimian, M., Nawara, D., Ibrahim, A., & Kashef, R. (2020).

Recommendation systems: Algorithms, challenges, metrics, and business opportunities. *Applied Sciences*, 10(21), 7748. MDPI.

<https://doi.org/10.3390/app10217748>

He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. S. (2017, April). Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web* (pp. 173-182).

<https://dl.acm.org/doi/abs/10.1145/3038912.3052569>

Jannach, D., & Adomavicius, G. (2016). Recommendation systems: Challenges, evaluation metrics, and business opportunities. *Communications of the ACM*, 59(11), 62-71. <https://doi.org/10.1145/2961031>

Koren, Y., et al. (2009). Matrix factorization techniques for recommender systems. *IEEE Computer Society*, 42(8), 30-37. <https://ieeexplore.ieee.org/document/5197422>

Nandi, B., Jana, S., & Das, K. P. (2023). Machine learning-based approaches for financial market prediction: A comprehensive review. *Journal of AppliedMath*, 1(2), 134. <https://doi.org/10.59400/jam.v1i2.134>

Liu, J., Dolan, P., & Pedersen, E. R. (2010). Personalized news recommendation based on click behavior. In *Proceedings of the 15th International Conference on Intelligent User Interfaces* (pp. 31-40). ACM. <https://doi.org/10.1145/1719970.1719976>

Rendle, S., Freudenthaler, C., Gantner, Z., & Schmidt-Thieme, L. (2012). BPR: Bayesian personalized ranking from implicit feedback. *arXiv preprint arXiv:1205.2618*. <https://arxiv.org/abs/1205.2618>

Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender systems: Introduction and challenges. In *Recommender Systems Handbook* (pp. 1-34). Springer.

https://doi.org/10.1007/978-1-4899-7637-6_1

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001, April). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web* (pp.

285-295). <https://dl.acm.org/doi/abs/10.1145/371920.372071>

Sharma, A., & Gera, S. (2013). A survey of recommendation system: Research challenges. *International Journal of Engineering Trends and Technology*, 4(5),

1989-1994. <https://www.ijettjournal.org>

Yousef, H. A. (2024). Recommender systems applications: Data sources, features, and challenges. *Information*, 15(10), 660. MDPI.

<https://doi.org/10.3390/info15100660>