

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INFORMÁTICA
SANTIAGO - CHILE



“ANÁLISIS Y PREDICCIÓN DEL ÉXITO DE UN VIDEOJUEGO MEDIANTE MINERÍA DE OPINIÓN EN RESEÑAS”

EMILIANO SANTANA

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN INFORMÁTICA

Profesor Guía: Liubov Dombrovskaja
Profesor Correferente: Roberto Asín Acha

Enero - 2024

DEDICATORIA

El trabajo realizado en la memoria es dedicado a la persona que mas paciencia me ha tenido en la vida y sobre todo en el periodo de realización del escrito: mi madre.

AGRADECIMIENTOS

La realización de esta memoria me ha hecho recordar cuanto apoyo he recibido a lo largo de la duración de la carrera y por ende veo apropiado agradecer a las personas que me han acompañado a través de esta.

A mi madre por siempre apoyarme y haberme criado.

A mis amigos los cuales me han acompañado incontables veces tanto fuera como dentro de la universidad y con los cuales se ha vuelto habitual convivir y se nota su falta cuando no hay interacción.

A mis tíos los cuales siempre están dispuestos a echarme una mano en caso de que lo necesite.

Y a mi profesora guía por la paciencia que me ha tenido y los consejos que me ha dado para seguir adelante.

Sin todos ellos posiblemente no habría avanzado lo suficiente por lo que solo puedo decir, muchas gracias.

RESUMEN

Resumen— La abundancia de la oferta de videojuegos en el mercado dificulta la búsqueda del juego para los jugadores y también obstaculiza el camino al éxito para los desarrolladores independientes, que no cuentan con medios de promoción de sus juegos y cuyos juegos pueden pasar desapercibidos. La existencia de plataformas de juegos en línea como STEAM y el alto número de reseñas en ellas pueden proporcionar información valiosa para predecir el éxito de un videojuego o incluso para predecir los aspectos más interesantes para los jugadores.

El presente trabajo emplea la minería de opinión para evaluar el éxito de un videojuego asignando a cada oración de una reseña una valoración positiva, neutra o negativa. A través de varias iteraciones, se logró un algoritmo no supervisado de extracción y clasificación de aspectos.

Para validar los resultados de esta clasificación, se contrastaron los números de aspectos positivos y negativos con el éxito de 25 juegos de 3 géneros distintos, logrando comprobar la existencias de correlación positiva con aspectos positivos y negativa con aspectos negativos. El factor de mayor correlación con el éxito del juego fue el ratio entre reseñas positivas y negativas en el mes anterior al ranking obtenido.

Palabras Clave— *Minería de Opinión; Videojuegos; Reseñas; Análisis de Sentimientos; Steam*

ABSTRACT

Abstract— The abundance of video game offerings on the market makes it difficult for gamers to find a game and also hinders the path to success for independent developers, who have no means of promoting their games and whose games may go unnoticed. The existence of online gaming platforms such as STEAM and the high number of reviews on them can provide valuable information to predict the success of a video game or even to predict the most interesting aspects for gamers.

The present work employs opinion mining to evaluate the success of a video game by

assigning each sentence of a review a positive, neutral or negative rating. Through several iterations, an unsupervised aspect extraction and classification algorithm was achieved. To validate the results of this classification, the numbers of positive and negative aspects were contrasted with the success of 25 games of 3 different genres, managing to verify the existence of positive correlation with positive aspects and negative correlation with negative aspects. The factor with the highest correlation with the success of the game was the ratio between positive and negative reviews in the month prior to the obtained ranking.

Keywords— Opinion Mining, Video games, Reviews, Sentiment Analysis, Steam

GLOSARIO

ABSA: *Aspect-Based Sentiment Analysis* o Análisis de Sentimientos Basado en Aspectos

API: *Application Programming Interface* o Interfaz de Programación de Aplicaciones

ASCII: *American Standard Code for Information Interchange* o Código Estándar estadounidense para el Intercambio de Información

APC: *Aspect Polarity Classification* o Clasificación de Polaridad de Aspectos

ATEPC: *Aspect Term Extraction & Polarity Classification* o Extracción de Términos de Aspectos y Clasificación de Polaridad

BERT: *Bidirectional Encoder Representations from Transformers* o Representación de Codificador Bidireccional de Transformadores

CNN: *Convolutional Neural Network* o Redes Neuronales Convolucionales

CUDA: *Compute Unified Device Architecture* o Arquitectura Unificada de Dispositivos de Cómputo

CRISP-DM: *Cross-Industry Standard Process for Data Mining*

CSV: *Comma-Separated Values* o Valores Separados por Comas

GPU: *Graphics Processing Unit* o Unidad de Procesamiento Gráfico

JSON: *JavaScript Object Notation*

LSTM: *Long Short-Term Memory* o Memoria de Corto Plazo Largo

MCDM: *Multiple-Criteria Decision Making* o Toma de Decisiones de Criterios Múltiples

ML: *Machine Learning* o Aprendizaje Automático

MMORPG: *Massively Multiplayer Online Role-Playing Game* o Videojuegos de Rol Multijugador Masivos en Línea

NLP: *Natural Language Processing* o Procesamiento del Lenguaje Natural

RNN: *Recurrent Neural Network* o Redes Neuronales Recurrentes

RPG: *Role-Playing Game* o Juego de Rol

SA: *Sentiment Analysis* o Análisis de Sentimientos

SVM: *Support Vector Machines* o Máquinas de Vectores de Soporte

ÍNDICE DE CONTENIDOS

RESUMEN	IV
ABSTRACT	IV
GLOSARIO	VI
ÍNDICE DE FIGURAS	IX
ÍNDICE DE TABLAS	IX
INTRODUCCIÓN	1
CAPÍTULO 1: DEFINICIÓN DEL PROBLEMA	3
1.1 Contexto	3
1.2 Objetivos	6
1.2.1 Objetivo General	6
1.2.2 Objetivos Específicos	6
CAPÍTULO 2: MARCO CONCEPTUAL	7
2.1 Minado de Opinión	7
2.1.1 Aprendizaje No Supervisado	8
2.1.2 <i>Cross-Domain SA</i>	8
2.1.3 <i>Fine-Grained SA</i>	9
2.1.4 <i>Aspect-Based SA</i>	9
2.1.5 <i>Deep Learning</i>	10
2.1.6 <i>Fine-Tuning</i>	10
2.1.7 Detección de Sarcasmo	11
2.2 Reseñas de videojuegos y la Minería de opinión	11
CAPÍTULO 3: PROPUESTA DE SOLUCIÓN	14
3.1 Metodología: CRISP-DM	14
3.2 Obtención de Reseñas	17
3.2.1 Formato	18
3.3 Iteraciones	19
3.3.1 Conocimiento de los datos	20
3.3.2 Modelo de grano fino	20
3.3.3 Afinación del modelo	21
3.3.4 Reestructuración	22
3.3.5 Aplicación en un periodo de tiempo	23
3.3.6 Variaciones	24
3.4 Algoritmo Resultante	24
3.4.1 Corpus	24
3.4.2 Obtención de los datos	26

3.4.3	Creación del <i>dataset</i> APC o corpus	27
3.4.4	Entrenamiento y Ejecución	27
CAPÍTULO 4: VALIDACIÓN DE LA SOLUCIÓN		28
4.1	Aprendizaje no supervisado y sus validaciones	28
4.2	Coeficiente de Correlación de Spearman	28
4.3	Coeficiente de Correlación de Rango de Kendall	29
4.4	Caso Base	30
4.5	Caso Fecha Filtrada	32
4.6	Caso Juegos Filtrados	34
4.7	Caso Juegos y Fecha Filtrados	35
4.8	Correlación Aspectos Postivos-Negativos	35
4.9	Comparación de coeficientes de correlación	36
CAPÍTULO 5: CONCLUSIONES		38
ANEXOS		41
REFERENCIAS BIBLIOGRÁFICAS		49

ÍNDICE DE FIGURAS

1	Plataforma <i>Steam</i>	3
2	Infografía de los ingresos de juegos independientes.	4
3	JSON de una reseña obtenida mediante (Valve Corporation, s.f.)	5
4	Esquema de flujo de CRISP-DM	16
5	Top 20 Aspectos Identificados de <i>Library of Ruina</i>	22
6	Top 20 Aspectos Identificados y Clasificados de <i>Baldur's Gate 3</i>	23
7	Explicación de la correlación negativa	29
8	Correlación entre el Top de Ventas y el Ratio Positivo-Negativo	36

ÍNDICE DE TABLAS

1	Plantilla para presentar y estructurar artículos siguiendo CRISP-DM	17
2	Reseñas pertenecientes al top 25	25
3	Reseñas Top 10 Aventura	25
4	Reseñas Top 10 RPG	26
5	Reseñas Top 10 Fantasía	26
6	Aspectos Identificados	31
7	Coeficiente de Correlación por Aspectos Identificados	31
8	Coeficiente de Correlación por ratio de Aspectos-Reseñas	32
9	Coeficiente de Correlación por Aspectos Identificados (Fecha Filtrada)	32
10	Aspectos Identificados (desde 1 de Septiembre)	33
11	Coeficiente de Correlación por ratio de Aspectos-Reseñas (Fecha Filtrada)	33
12	Coeficiente de Correlación por Aspectos Identificados (Juegos Filtrados)	34
13	Coeficiente de Correlación por ratio de Aspectos-Reseñas (Juegos Filtrados)	34
14	Coeficiente de Correlación por Aspectos Identificados (Juegos y Fecha Filtrados)	35
15	Coeficiente de Correlación por ratio de Aspectos-Reseñas (Juegos y Fecha Filtrados)	35
16	Coeficiente de Spearman por ratio de aspectos Positivos-Negativos	36

INTRODUCCIÓN

El desarrollo de videojuegos se ha vuelto mucho mas accesible con los avances tecnológicos, lo que ha producido un gran crecimiento en el mercado de los videojuegos, pero también ha provocado su saturación. Como consecuencia se pueden encontrar múltiples juegos de nicho para los distintos gustos, pero no todo juego llega a lograr un ansiado éxito.

Esto se denota más cuando un videojuego no logra adquirir atención suficiente para recibir una retroalimentación, la cual se manifiesta en forma de reseñas de los usuarios y contiene diversas informaciones como los puntos fuertes o errores cometidos referentes al juego. Permitiendo a sus desarrolladores lograr mejoras y mantenerse a flote en la competencia.

El trabajo presentado a continuación busca observar y analizar las reseñas de los videojuegos mediante la minería de opinión, con el objetivo de identificar una relación tangible frente al éxito de los videojuegos, especialmente en la plataforma *Steam*, la cual se ha convertido en en principal mercado de los juegos para computadores de escritorio, y además contando con la disponibilidad de reseñar videojuegos, para luego realizar solicitudes de obtención de las reseñas como datos manipulables.

Frente a la problemática se utiliza la metodología *CRISP-DM* para mejorar y ajustar continuamente el modelo y/o algoritmo, la cual se basa en iteraciones constantes sobre el trabajo realizado. Se utilizan métodos estadísticos no paramétricos como lo son el coeficiente de correlación de Spearman y de Kendall para validación.

El siguiente documento se estructura de la siguiente manera;

- **Capítulo 1 - Introducción:** Se explica con mayor detalle la situación existente en el mercado de videojuegos, el desarrollo de estos junto a los problemas que involucran, además de los objetivos del escrito.
- **Capítulo 2 - Marco Conceptual:** Se presentan distintos casos de estudios del Minado de Opinión y su aplicación respecto a reseñas de videojuegos, luego se explican varios términos relacionados al escrito.
- **Capítulo 3 - Propuesta de Solución:** Mediante el uso de *CRISP-DM* se procede a realizar un algoritmo de creación de conjuntos de datos empleando tres clasificaciones de ventas de distintos géneros de videojuegos, el *dataset* es utilizado en el entrenamiento de otro algoritmo de extracción de aspectos a reseñas pertenecientes a una clasificación de ventas correspondiente a los tres géneros juntos, ambos algoritmos utilizan métodos no supervisados de análisis de sentimientos.
- **Capítulo 4 - Validación de la Solución:** Mediante el uso de los coeficientes de correlación ρ de Spearman y τ de Kendall se busca analizar e identificar una posible correlación entre la clasificación de ventas y los aspectos identificados por el algoritmo, los cuales se identifican con su respectiva polaridad de sentimiento.

- **Capítulo 5 - Conclusión:** Abarca el aprendizaje realizado y el resultado a lo que llega, además se proponen diferentes ideas propuestas para un futuro trabajo retomado.

CAPÍTULO 1

DEFINICIÓN DEL PROBLEMA

1.1. Contexto

En los años recientes la tecnología ha avanzado a pasos agigantados al punto de que no solo se ha vuelto una herramienta de trabajo, sino que también un método de entretenimiento al alcance de la mayoría de la población. Una de las formas de entretenimiento más comunes son los videojuegos, un mercado que se ha logrado expandir muy rápidamente en los últimos años al punto que se ha vuelto fácil encontrar juegos con géneros, temáticas, o características de interés específico para el usuario.

Mientras que el crear un videojuego desde cero se ha simplificado con la cantidad de herramientas disponibles, como por ejemplo, el motor de videojuegos *Unity* en el cual una vez desarrollado un producto se puede exportar para ser compatible con varios tipos de dispositivos como consolas, computadores, o teléfonos, para luego ser comercializado por ejemplo en *Steam* la cual es la plataforma líder de *e-commerce* respecto a videojuegos y donde un desarrollador pagando una cuota de 100 dólares ya puede publicar su obra como se observa en (Valve Corporation, 2017).

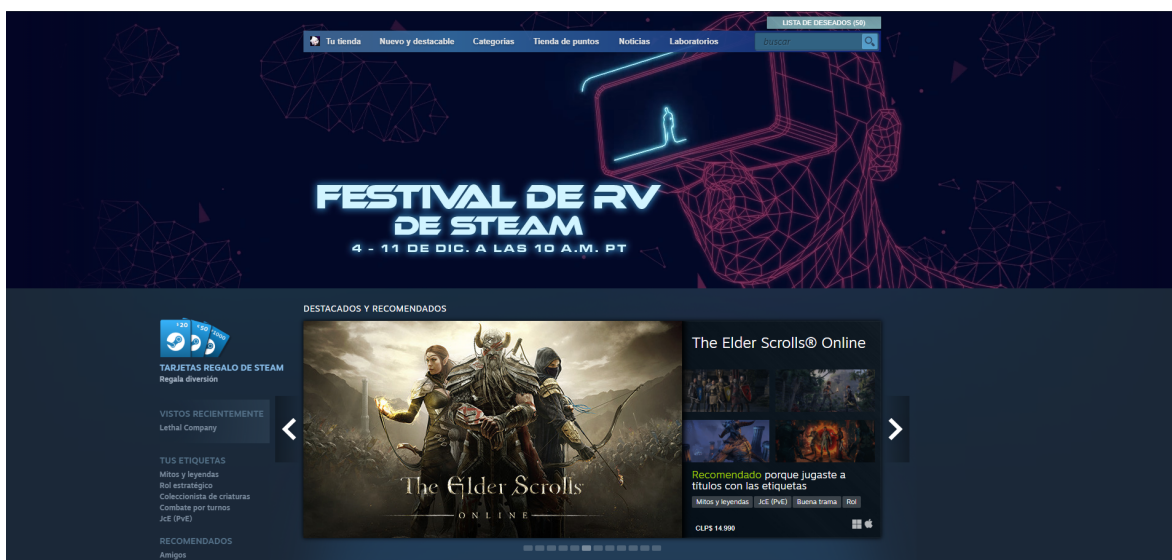


Figura 1: Plataforma *Steam*.
(«Steam Store», s.f.)

Pero esto también ha significado que el mercado se ha saturado de oferta con una gran competencia, en donde el destacar se ha vuelto un desafío para los desarrolladores independientes al punto que muchos apenas obtienen ganancias (Véase la figura 2), y esto además provoca que una vez se produce un videojuego con un éxito relativo dentro de un nicho es-

pecífico, el desarrollador tienda a convertirlo en su área de confort con tal de ser capaz de seguir realizando publicaciones.

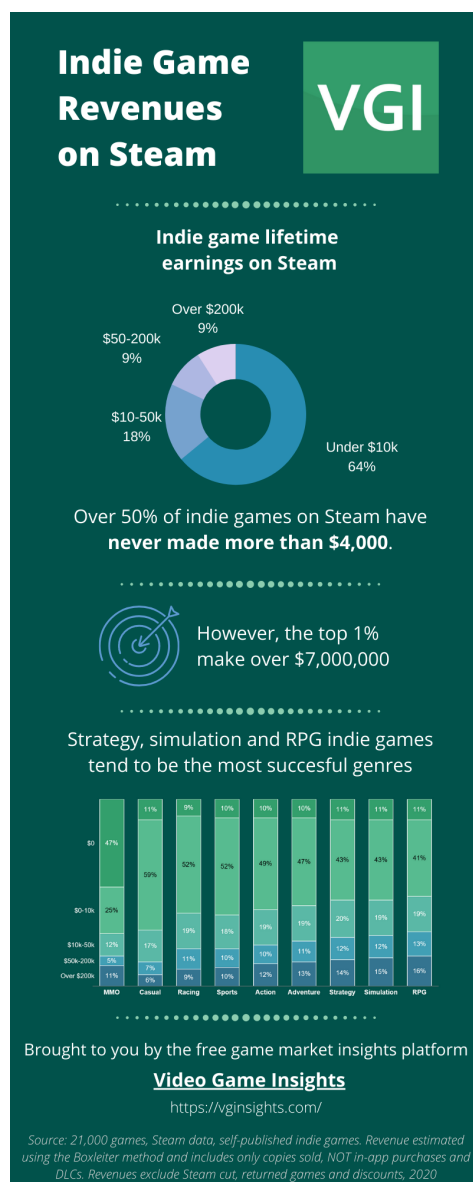


Figura 2: Infografía de los ingresos de juegos independientes.
 (Video Game Insights, 2020).

Debido a la masificación del internet, con unos pocos clics se puede cambiar de contenido en cuanto el actual pierda interés para el usuario, aunque en el caso de los videojuegos al pasar por un momento de popularidad se puede obtener una gran ganancia no solo monetaria sino también informativa, debido a la retroalimentación obtenida en el caso de que los jugadores dejen reseñas.

Las reseñas por su parte son una gran fuente de información especialmente al considerar

```

{
  "cursor": "AoIIPwYYanmY3tcD",
  "query_summary": {
    "num_reviews": 1,
    "review_score": 8,
    "review_score_desc": "Very Positive",
    "total_negative": 324,
    "total_positive": 5113,
    "total_reviews": 5437
  },
  "reviews": [
    {
      "author": {
        "last_played": 1489400332,
        "num_games_owned": 253,
        "num_reviews": 15,
        "playtime_at_review": 8457,
        "playtime_forever": 8457,
        "playtime_last_two_weeks": 0,
        "steamid": "76561197972393397"
      },
      "comment_count": 0,
      "hidden_in_steam_china": true,
      "language": "english",
      "received_for_free": false,
      "recommendationid": "123662729",
      "review": "a game about pudding\n10/10 neps",
      "steam_china_location": "",
      "steam_purchase": true,
      "timestamp_created": 1665453568,
      "timestamp_updated": 1665453568,
      "voted_up": true,
      "votes_funny": 0,
      "votes_up": 1,
      "weighted_vote_score": "0.523809552192687988",
      "written_during_early_access": false
    }
  ],
  "success": 1
}

```

Figura 3: JSON de una reseña obtenida mediante (Valve Corporation, s.f.)

que, pese a la brevedad de un comentario, si se consideran una cantidad significativa de reseñas los datos que pueden derivar de estas pueden otorgar un margen de mejora importante al permitir que el desarrollador corrija sus falencias, mejore sus fortalezas, y entienda las preferencias de su público.

Por ejemplo, en el caso de adquirir un producto mediante *Steam*, el usuario tiene la capacidad de reseñar este, y pese a que el porcentaje de usuarios que realmente lo hace es menor al 2 % en promedio¹, al saber la utilidad posible de estas opiniones, el mismo sitio coloca a disposición una API con la cual se pueden obtener éstas en formato JSON solo requiriendo el identificador del videojuego, retornando lo observable en la figura 3, contando adicionalmente con una disposición de datos útiles tales como: los números generales referentes a la aprobación de los usuarios, datos del usuario que hizo la reseña, la marca de tiempo de la

¹Según (Birkett, 2018) el ratio entre reseñas y ventas en promedio es 1 a 82 respectivamente

publicación, cuantos usuarios la encontraron útil o divertida, entre otros.

Basándose en lo anterior nace la propuesta de utilizar estos datos a disposición mediante el minado de opinión para lograr identificar qué factores influyen en el éxito de un videojuego, con este objetivo en mente se distinguen dos necesidades importantes a definir siendo el primero el origen y estructura de los datos debido a la cantidad adicional de información inesperada que pueden proveer además del comentario de texto, mientras que el segundo es la técnica a ocupar para el análisis de la información considerando que el campo del también llamado Análisis de Sentimientos está continuamente en desarrollo. Una vez realizadas la selección, el abarque cambia dependiendo de si el videojuego ya ha sido publicado o no pues influye en la obtención de las reseñas, el caso prístino considera la disposición de opiniones dirigidas a la obra examinada tal que permitan el estudio de la relación causa-efecto asociada al éxito del producto con el objetivo de entender él porque es popular, mientras que el caso suplementario trata las similitudes designadas entre el videojuego y sus referentes designados para primero procesar estos últimos a fin de identificar la relevancia de las características mencionadas para deducir si es que la obra podrá ser popular.

1.2. Objetivos

1.2.1. Objetivo General

Identificar los factores de influencia en el éxito de un videojuego en base al análisis de sus reseñas y sus variaciones en el tiempo mediante el Minado de Opinión.

1.2.2. Objetivos Específicos

- Permitir estimar la valorización futura de un videojuego en base a la identificación de las similitudes tanto positivas como negativas respecto a otros videojuegos referentes.
- Determinar las razones que influyen en la definición de la opinión de los jugadores a lo largo de la vida del videojuego y sus actualizaciones.
- Comparar las fuentes y conformación de reseñas disponibles junto con las técnicas de minado de opinión para identificar, estructurar, y/o combinar las más apropiadas al contexto de la finalidad.

CAPÍTULO 2

MARCO CONCEPTUAL

Para un mejor entendimiento de los distintos conceptos a utilizar en el proceso de desarrollo, el presentar variedad de definiciones respecto al análisis de sentimientos se vuelve fundamental para el entendimiento de los futuros pasos, además seguido de distintos trabajos académicos relacionados a los videojuegos para contextualizar sobre la relación entre ellos.

2.1. Minado de Opinión

El también llamado Análisis de Sentimientos, *Sentiment Analysis* (SA), consiste en el uso de la Lingüística Computacional para lograr el Procesamiento del Lenguaje Natural (NLP), la interacción entre el lenguaje humano, con el objetivo de identificar la intención o sentimiento del autor al emitir el mensaje.

Existen dos aproximaciones principales para el manejo de los datos siendo los métodos basados en léxico y en *machine learning* o aprendizaje automático, los cuales también se mezclan para crear métodos híbridos.

- **Basado en Léxico:** Esta aproximación consta de la utilización de reglas y diccionarios de palabras las cuales cuentan con un puntaje para indicar la polaridad del sentimiento, con la cual se obtiene una ponderación de la polaridad de la frase, para así lograr asignar un puntaje y por ende un sentimiento al texto.
- **Machine Learning:** El método basado en *Machine Learning* (ml) busca automatizar la predicción de los sentimientos al entrenar un clasificador, contando con la ventaja de ser capaz de procesar grandes cantidades de datos y aprender de estos para generar un resultado de mayor precisión, aunque esto también se vuelve su desventaja en caso de contar con un conjunto de datos pequeño u no adecuado.

Dentro del ámbito existen tres paradigmas principales utilizados para clasificar el modo de aprendizaje:

- **Supervisado:** Implica el uso de datos totalmente estructurados y etiquetados.
- **No Supervisado:** Radica en la inferencia de similitudes entre los datos sin etiquetar.
- **Aprendizaje Reforzado:** Utiliza agentes inteligentes los cuales buscan maximizar una recompensa acumulada obtenida por la variación de sus acciones.

Además de los anteriores también existen otras tales como:

- **Semi-Supervisado:** Consiste en la utilización tanto de datos etiquetados como sin etiquetar.
- **Aprendizaje Profundo:** Utiliza algoritmos con altos grados de abstracción computacional para identificar relaciones.

2.1.1. Aprendizaje No Supervisado

El Aprendizaje Automático No Supervisado comprende los datos como variables aleatorias y los modelos buscan relaciones o características de forma autónoma, simplificando la necesidad de obtener con antelación datos etiquetados los cuales de no estar corrector provocarían ruido en el modelo.

(Punetha & Jain, 2023b) indaga en un acercamiento a la teoría de juegos mediante los juegos bayesianos, la cual es una teoría de juegos con información incompleta de las características de otros jugadores, utilizada para la creación de un método no supervisado, logrando resultados con una exactitud y precisión del 94 % y 95 % respectivamente.

Otro método no supervisado interesante es el implementado mediante el uso de reglas difusas, siendo parte de la lógica difusa y valiéndose de valores entre 0 y 1 en contraste a una lógica binaria, este se observa en (Vashishtha & Susan, 2019) y es adaptable a cualquier léxico.

2.1.2. Cross-Domain SA

Debido a que no siempre existen datos disponibles, surge el *Cross Domain Sentiment Analysis*, el cual afronta el problema de adaptar un dominio de origen cuyo tema difiere del de los datos objetivos. La utilización de modelos ya entrenados en dominios ajenos es una de las varias formas de ahorrar costos, ya sean de datos o tiempo debido a la cantidades necesarias de estos involucradas en los entrenamientos, y que dependiendo del tipo de adaptación se distinguen por la cantidad de dominios de origen utilizados.

(Kotelnikova et al., 2023) menciona dos conclusiones importantes a considerar durante la implementación del *Cross-Domain SA*, el uso añadido de corpus o conjunto de datos utilizados, siendo estos ajenos al dominio no mejora el rendimiento a no ser que el conjunto de datos del dominio sea extremadamente bajo, y también utilizar conjuntos de corpus ajenos como reemplazo no es recomendable, salvo en el caso de ausencia de datos del dominio o que estos sean en muy baja cantidad.

(Chen & Wan, 2022) crea una técnica no supervisada para lograr la adaptación de dominio y alcanzar mejores resultados para la *Aspect-Based SA* y la extracción de aspectos, mediante la utilización de la técnica de maximización de información mutua, con esto último siendo una dependencia estadística entre variables.

2.1.3. *Fine-Grained SA*

El concepto de grano fino o *Fine-Grained Sentiment Analysis* se refiere a un entendimiento de los datos analizados a un nivel de oración o característica en contraste de la ocupación de una reseña completa siendo este llamado *Coarsed-Grained Sentiment Analysis* o de grano grueso. Como explica (Guo et al., 2021), lo anterior se considera en los niveles de documento, oración, y aspecto, siendo esto último estrechamente ligado con análisis de sentimientos basado en aspectos.

La aproximación híbrida basada en reglas realizada por (Zainuddin et al., 2018) ocupa las oraciones identificando si son objetivas o subjetivas para la realización de un análisis de sentimientos de grano fino, en caso de ser subjetivas realiza un análisis adicional para determinar su polaridad.

2.1.4. *Aspect-Based SA*

ABSA o Análisis de Sentimientos Basado en Aspectos, es un método que ocupa la categorización de conceptos para obtener mediante *Machine Learning* los sentimientos respecto a estos, siendo un método híbrido al también ocupar ideas de la aproximación basada en léxico como la utilización de reglas o diccionarios pero ligados a una categoría en especial, obteniendo así un resultado más contextual y específico del tema. Existen dos sub-tareas principales para este análisis de sentimientos, la clasificación de polaridad de aspectos (APC), y la extracción de términos de aspectos (ATE).

(Punetha & Jain, 2023a) combina la teoría de juegos y la Toma de Decisiones de Criterios Múltiples o MCDM para la creación de dos *frameworks* de análisis de sentimientos basado en aspectos, llegando a una exactitud del 91 % se considera una alternativa con futuro a la utilización de otros modelos de ML o basados en léxico. Lamentablemente, cuenta con una debilidad en su implementación actual, fallando en situaciones donde existen palabras que no están en su léxico de sentimiento.

(Yang et al., 2021) propone un modelo llamado LCF-ATEPC el cual está orientado principalmente al idioma chino. Este modelo multitarea es capaz de realizar ambas sub-tareas relacionadas a ABSA, la clasificación de polaridad y la extracción de aspectos.

Una implementación híbrida utilizando *long short-term memory* (memoria larga de corto plazo), y corte de grafos para n-grama se presenta en (Nandi et al., 2023), la cual sirve para

la tarea de clasificación de polaridad y es compatible para combinarse con otros métodos.

2.1.5. *Deep Learning*

El Aprendizaje Profundo o *Deep Learning* como se ha mencionado anteriormente se basa en el uso de algoritmos con grados de abstracción computacional altos. Entre los métodos principales, las redes neuronales convolucionales (*CNN*), y las redes neuronales recurrentes (*RNN*), siendo las primeras basadas en el uso de capas ocultas y filtros, mientras las *RNN* utilizan sus resultados para una retroalimentación.

(Oh et al., 2021) propone su *Deep Contextualized Relation-Aware Network*, el cual es capaz de obtener los aspectos, la opinión respecto a estos, y la polaridad referente al aspecto, siendo capaz de identificar múltiples aspectos al mismo tiempo o de forma singular.

El idioma indonesio presenta su propio modelo con (Kusumawardani et al., 2018), su objetivo es la normalización de texto obtenido por medios sociales, adaptando símbolos y palabras al lenguaje formal escrito, por lo que usa la incrustación de palabras de *Word2Vec*, logrando una exactitud de casi 80 % al normalizar mensajes enteros.

Otro lenguaje que también presenta un modelo dedicado es el urdu, en (Sehar et al., 2021) introduce un *framework* combinación de *CNN* y *LSTM* capaz de realizar análisis de sentimiento no solo para texto, también para contenido multimedia.

2.1.6. *Fine-Tuning*

El concepto de *Fine-Tuning* o Ajuste Fino surge como una alternativa para reusar o transferir pesos de un modelo ya entrenado y ejecutar un nuevo entrenamiento con otros datos. Un ejemplo común es la utilización de modelos entrenados con grandes corpus generales, para luego agregar capas mediante entrenamiento con un corpus específico al dominio que se necesita aplicar.

La adición de un número relativamente pequeño de textos ajenos al dominio puede facilitar la realización del *Fine-Tuning* como explica (Kotelnikova et al., 2023).

(Li et al., 2021) abarca los mejores valores para los hiperparámetros de *BERT* cuando se le realiza *fine-tuning*, indicando como mejores valores una tasa de aprendizaje de $2e-5$, y un tamaño de lotes de 32, para luego proceder a mejorar el ajuste mediante una estrategia de capa congelada dinámica.

2.1.7. Detección de Sarcasmo

La Detección de Sarcasmo se ha vuelto un problema al momento de analizar información recopilada de distintos medios en los que se puede emitir comentarios libremente como las redes sociales. Por lo tanto se ha vuelto un desafío a solucionar mediante el Minado de Opinión, ya que aproximaciones como el basado en léxico es incapaz de procesarlo correctamente, pero ocupando *Machine Learning* como se observa en (Bhakuni et al., 2022) hay clasificadores capaces de llegar a más de un 90 % de precisión.

(Shangipour ataei et al., 2020) realiza una investigación del sarcasmo en corpus pertenecientes a *Twitter* y *Reddit* donde concluye la necesidad de utilizar a futuro métodos contextuales e incrustación de palabras, junto a la normalización de texto obtenido por medios sociales.

Los medios sociales no solo incluyen sarcasmo, sino también el uso de emoticones, y (Gupta et al., 2020) presenta un *framework* capaz de identificar los casos y asignar polaridades a estos, mediante el uso de las técnicas de redes neuronales artificiales y árbol de clasificación y regresión.

Un estudio realizado por (Verma et al., 2021) presenta una comparación entre las distintas técnicas de detección de sarcasmo, en el que destaca *sAtt-BiLSTM-ConvNet* debido a sus prominentes resultados con el *dataset SemEval 2015*, llegando a obtener un 97 % de certeza en su predicción.

Un modelo dedicado a identificar el sarcasmo en el idioma árabe, el cual es uno de los mas complicados, es presentado en (Nayel et al., 2021) utilizando como base *SVM* o *Support Vector Machine* con un rendimiento del 85 % superando otros algoritmos y logrando una exactitud del 74 % en una prueba a ciegas.

(Touahri & Mazroui, 2021) implementa un método de detección de sarcasmo capaz de procesar múltiples idiomas debido a la disponibilidad de corpus etiquetado, la construcción automática de léxico, y la identificación de aspectos sarcásticos, llegando a un 89 % de certeza.

Otra aplicación existe en (Cortis & Davis, 2022), esta vez contando con el idioma inglés y maltés, se divide la prueba en cinco factores: subjetividad, polaridad, emoción, ironía, y sarcasmo, junto a un calculo adicional de la predicción del lenguaje para observar la efectividad del modelo inglés-maltés.

2.2. Reseñas de videojuegos y la Minería de opinión

La minería de opinión o también llamado análisis de sentimientos es un campo en constante desarrollo en el cual se han realizado varios estudios para abarcar su potencial y limitantes actuales, como en el caso del análisis de las reseñas de videojuegos en donde se pueden encontrar distintos acercamientos al tema.

(Urriza & Clarino, 2021) utiliza una clasificación de dos niveles utilizando la clasificación basada en características y la clasificación de polaridad en la cual resulta similar a la técnica binaria pero adicionalmente considerando un estado neutral. Pese a que realiza las clasificaciones de forma independiente es importante mencionar que los datos obtenidos permiten un entendimiento mayor al clarificar que aspectos específicos contienen una mayor o menor aprobación.

Aunque respecto a múltiples clasificaciones, (Chakraborty et al., 2018) propone un clasificador por votación, el cual utiliza siete clasificadores provenientes de cuatro algoritmos diferentes para asignar una polaridad a comentarios emitidos en la plataforma *Twitter* junto con una valuación entre 1 y 5 generada a partir de las votaciones positivas obtenidas internamente por cada reseña.

En (Eberhard et al., 2018) se busca indagar el beneficio de la “utilidad” de las reseñas obtenidas desde la plataforma *Steam* antes mencionada considerando que estas cuentan con un campo numérico indicando la cantidad de usuarios a los que les pareció útil el comentario. Algunas de las conclusiones que se desprenden del escrito apuntan a que a más “útil” el comentario tiende a ser más extenso, crítico, y detallista, mientras que se puede dar el caso donde sean opiniones cortas o sin sentido y aun así ser “útiles” debido al contexto del autor, el humor, o un origen externo como un video.

Al mencionar *Steam* como una fuente para la obtención de reseñas es importante analizar la información empíricamente como se realiza en (Lin et al., 2019), en el cual realizan hallazgos tales como que los jugadores tienen un tiempo promedio de 13.5 horas al momento de realizar sus reseñas, que los tiempos de juego de las reseñas negativas tienden a ser la mitad que de las positivas, en el caso de los juegos *Free-to-Play* un gran número de reseñas se realiza tan solo a la hora de juego, y el como diferentes de tipos de videojuegos también implica distintos tipos de reseñas.

De la misma forma (Andreev et al., 2021) ocupa las críticas de *Steam* analizando si fueron positivas o negativas en contraste con los sentimientos que verdaderamente representan y la magnitud de que tan presentes están en el comentario independiente de la polaridad, resultando en que los sentimientos si tienden a alinearse con el carácter positivo o negativo del escrito mientras que en relación a la magnitud observan que en promedio es mayor cuando la reseña es negativa, pero en ambos casos sus medias son iguales dando a entender que la compensación de las opiniones largas respecto a las cortas es mayor en caso de que sean negativas.

Otro tipo de análisis a las reseñas de *Steam* es realizado por (Kang, 2017), donde se investigan características asociadas al usuario respecto al valor de utilidad de la reseña.

Otra plataforma utilizada para la extracción de datos es *Youtube* como se logra observar en (Santos et al., 2021) donde analiza el videojuego *Just Dance* en el cual el jugador imita una coreografía, a fin de identificar el efecto que presenta en la salud de este mediante un alcance en tres dimensiones: usabilidad, experiencia de usuario, y percepción de calidad de vida.

Metacritic es la plataforma que utiliza (Kasper et al., 2019) para la extracción de 319 mil reseñas que mediante análisis correlacional busca una relación frente a la puntuación de la reseña. Un punto importante a destacar es la mención del género del videojuego como un factor determinante y como se demuestra por análisis predictivos y correlacionales.

Nuevamente utilizando *Metacritic*, (Ruseti et al., 2020) analiza 200 mil reseñas en una *pipeline* escalable mediante tres modelos distintos: *SVM*, *Deep Neural Networks*, y *Naïve-Bayes* multinomial, pero con resultados los cuales no logran superar una precisión del 67 %.

Con la misma plataforma anterior, (Strååt & Verhagen, 2017) utiliza dos sagas de videojuegos, *Mass Effect* y *Dragon Age*, y en el análisis realizado menciona que con cada nuevo lanzamiento aumentan las reseñas del videojuego mas reciente, pero que ese incremento se debe principalmente a reseñas negativas.

(Viggiato et al., 2022) enfoca en los límites buscando la razón, por el cual se producen los errores de clasificaciones en las reseñas de videojuegos, con lo que identifica necesidades tales como: estratificar los datos por género del videojuego porque las terminologías de lo positivo o negativo pueden variar dependiendo de esto, el manejo de técnicas capaces de procesar textos que incluyan ventajas y desventajas para no confundir a los clasificadores, el procesamiento de comparaciones entre videojuegos o versiones anteriores, y resolver la forma de manejar el sarcasmo al analizar reseñas que lo incluyan.

Un análisis fuera de lo normal se realiza por (Guzsvinecz, 2023) ya que investiga un género específico de videojuegos llamado "*Souls-Like*", los cuales se caracterizan en general por su alta dificultad, encontrando una ligera correlación entre las reseñas positivas y el tiempo de juego, e identificando algunos factores con mayor tendencia a tener reseñas positivas.

Por otro lado, (Kewirata et al., 2022) realiza un análisis de sentimientos a datos obtenidos mediante una encuesta para identificar el comportamiento de los usuarios, como sus tendencias a gastar dinero y tiempo en ciertos casos.

Un uso extremadamente útil para casos de *testing* se presenta en (Sykownik et al., 2019), ya que utilizando la *pipeline* implementable en *Unity* permiten la conversión de la voz del usuario a datos textuales generados durante el periodo de juego, a los cuales se aplica análisis de sentimientos, y muestra los datos obtenidos en una aplicación web.

(ARIK, 2022) utiliza reseñas extraídas desde *Reddit* para videojuegos *MMORPG* con el objetivo de analizar que los hace tan populares, y mencionando como principal fuerte la comunidad que se genera entorno a estos.

CAPÍTULO 3

PROPUESTA DE SOLUCIÓN

Es importante resaltar que el minado de opinión se vuelve un método de análisis de suma importancia al considerar la realización de un estudio que no nace de estadísticas ya entregadas, sino que busca indagar en una relación implícita, obtenida por un procesamiento abstracto correspondiente como podría ser una ejecución de redes neuronales.

Al utilizar lo anterior para la creación de la propuesta, en vista de la necesidad de aplicar una metodología para estructurar el proceso de trabajo se realiza una selección de esta, para continuar con la explicación de la obtención de los datos y el formato en el que se adquieren. Luego, se realizan las distintas iteraciones aplicando la metodología escogida, obteniendo así un algoritmo con el que se busca solucionar el problema anteriormente planteado.

3.1. Metodología: CRISP-DM

Cross Industry Standard Process for Data Mining, o simplemente *CRISP-DM* es una metodología altamente utilizada para la Minería de Datos, esto debido al ser un método iterativo mantiene una gran compatibilidad con trabajos cuyo flujo pese a ser estructurado no es absolutamente lineal. El flujo general es observable en la figura 4 aunque en caso de ser necesario se puede retroceder de fase, otorgando así una mayor flexibilidad al método como se indica en (Sngular, 2016).

Esta metodología consiste de 6 pasos:

1. Comprensión del Negocio

El primer paso consiste en analizar y definir los objetivos del proyecto con tal de identificar el como se trabajará y en que momento se ha sido exitoso. Al ser el paso más conceptual también es el que requiere una mayor creatividad debido a que estructura las futuras tareas a realizarse, por lo que el tiempo dedicado a este paso adquiere una importancia vital.

2. Comprensión de los Datos

Considera tanto la obtención de los datos como el trabajo inicial hacia estos para entender su viabilidad, calidad, atributos, y posibles relaciones no detectadas anteriormente.

3. Preparación de los Datos

El tercer paso se puede sintetizar como la transformación de los datos obtenidos en bruto a su forma utilizable. Esto incluye tareas de limpieza y selección de atributos, las cuales han de realizarse en relación con criterios de filtrado para lograr una mejor calidad de los datos.

4. Fase de Modelado

Consiste de la selección y utilización del modelo a ocupar, además de definir sus parámetros óptimos, separar sus datos para el entrenamiento, y comprobar su precisión mediante un caso de prueba.

5. Evaluación

Este paso evalúa los resultados del modelo frente a los objetivos definidos en la Comprensión de los Datos, considerando una interpretación en profundidad para observar posibles irregularidades, conclusiones erróneas, u otros factores que afecten la factibilidad del modelo. Una revisión a los procesos realizados con anterioridad también debería de incluirse a fin de buscar posibles mejoras.

6. Implementación

La fase final varía dependiendo de la finalidad de la utilización del modelo, ya que esta puede ser una realización de un reporte, la creación de un software con su respectivo mantenimiento, o un proceso constante y automatizado para analizar información.

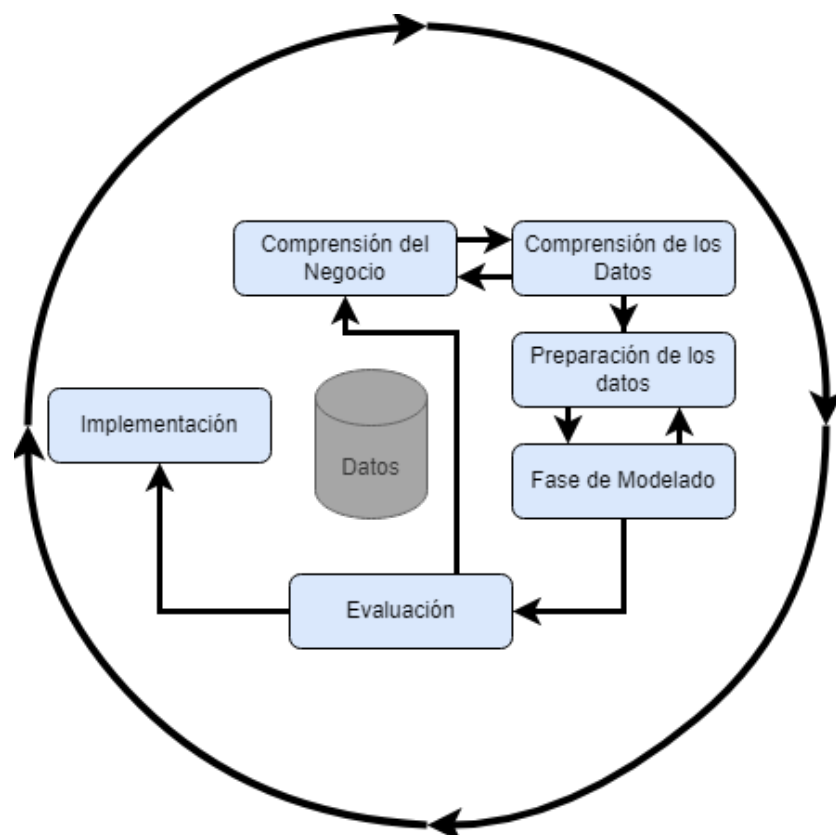


Figura 4: Esquema de flujo de CRISP-DM
(Dylan, 2019)

Adicionalmente, es relevante aludir que pese a la longevidad de la metodología, aún es altamente relevante como se menciona en (Schröer et al., 2021), donde se realiza un análisis de su utilización en 24 estudios con tal de verificar el porqué sigue siendo utilizada, cómo está siendo implementada, y en qué ámbitos tiene márgenes de mejora, además de incluir una estructuración simple de como incluir la metodología en un artículo científico como se logra observar en la tabla 1.

Con el entendimiento de los objetivos de esta memoria y los campos a trabajar se produce una necesidad de utilizar una metodología, la cual sea tanto compatible con la obtención y manejo de grandes cantidades de datos, como con tareas a realizarse mediante ensayo y error, por lo que al considerar tanto el flujo como la flexibilidad permitida por CRISP-DM se vuelve una solución idónea a los anteriores requisitos.

Tabla 1: Plantilla para presentar y estructurar artículos siguiendo CRISP-DM (Schröer et al., 2021)

Fase CRISP-DM	Metodos y aproximaciones como subsecciones
1. Comprensión del Negocio	1.1 Descripción textual en sección propia 1.2 Definir explícitamente el objetivo de la minería de datos
2. Comprensión de los Datos	2.1 Mención del origen de los datos y el proceso de recolección 2.2 Descripción estructural (modelo de datos, datos de ejemplo) 2.3 Estadísticas descriptivas obligatorias
3. Preparación de los Datos	3.1 Describir datos de entrada y salida 3.2 Métodos y aproximaciones (transformación, selección, limpieza)
4. Fase de Modelado	4.1 Mención del enfoque del modelo 4.2 Al menos la tecnología usada debería ser mencionada aquí 4.3 Creación del conjunto de prueba y entrenamiento
5. Evaluación	5.1 Definir métricas 5.2 Visualización del modelo y métricas
6. Implementación	6.1 Si la implementación está en el alcance, esta debe describirse

3.2. Obtención de Reseñas

Al considerar la utilización del Minado de Opinión se resalta la necesidad de una fuente de datos de tamaño considerable, estructurada, y de acceso público, dando como resultado al uso de una plataforma en específico: *Steam*.

Anteriormente se ha mencionado reiteradamente a la plataforma *Steam* en el escrito, contando con 20 años de experiencia desde su inauguración en 2003, nace como forma de actualizar los videojuegos de su autor *Valve Corporation* y luego permitiendo a terceros la inclusión de sus productos para distribuirlos, contando así como una tienda y una biblioteca digital. Entre los servicios adicionales que ofrece se encuentran *Steamworks* como interfaz de programación o *Workshop* para la distribución de contenido creado por los usuarios. Su utilización no solo se debe a su posición como líder del mercado de los videojuegos, sino también por su estructurados métodos disponibles en su servicio *Steamworks* para ser llamados principalmente por desarrolladores, con estos teniendo varios métodos exclusivos para procedimientos e integraciones a sus videojuegos publicados. Adicionalmente, su elección se vuelve prioritaria al notar la documentación disponible, lo cual facilita su implementación.

La API de (Valve Corporation, s.f.) permite los siguientes parámetros al ser llamada:

- **filter:** Tipo de obtención de las reseñas, ordenando por las creadas recientemente, actualizadas, o por utilidad, siendo este ultimo por defecto.
- **language:** El idioma de origen de las reseñas en caso de requerir uno en específico.
- **day_range:** Un número entero de hasta hace cuantos días atrás desde la solicitud se obtienen las reseñas en caso de solicitar por utilidad.
- **cursor:** Código de referencia para la obtención del lote siguiente de reseñas.
- **review_type:** La polaridad de la reseña, para abarcar solo positivas, solo negativas, o ambas.
- **purchase_type:** El tipo de adquisición del producto, por defecto se obtienen solo reseñas de productos comprados en la misma tienda, siendo posible adquirir productos mediante códigos obtenidos externamente y canjearlos, o ambos.
- **num_per_page:** La cantidad de reseñas obtenidas, por defecto siendo 20 con un máximo de 100.
- **filter_offtopic_activity:** Por defecto se excluyen reseñas producidas en masa con objetivos malintencionados, también llamado *Review Bombing*.

3.2.1. Formato

Como resultado de la llamada se los obtienen datos en formato *JSON*:

- **success:** 1 en caso de respuesta exitosa.
- **query_summary:** Estadísticas generales del videojuego.
 - **num_reviews:** Cantidad de reseñas retornadas.
 - **review_score:** El puntaje de reseña del videojuego.
 - **review_score_desc:** Descripción textual del puntaje.
 - **total_positive:** Cantidad de reseñas positivas.
 - **total_negative:** Cantidad de reseñas negativas.
 - **total_reviews:** Cantidad de reseñas totales.
- **cursor:** Identificador para obtener el siguiente lote de reseñas.
- **reviews:** Las reseñas retornadas.
 - **recommendationid:** Identificador único de la reseña.
 - **author:** Datos del autor de la reseña.

- `steamid`: Identificador de *Steam* del usuario
- `num_games_owned`: Cantidad de juegos pertenecientes al usuario.
- `num_reviews`: Numero de reseñas realizadas por el usuario.
- `playtime_forever`: Cantidad total de tiempo en la aplicación.
- `playtime_last_two_weeks`: Cantidad de tiempo en la aplicación durante las ultimas dos semanas.
- `playtime_at_review`: Cantidad de tiempo en la aplicación la momento de realizar la reseña.
- `last_played`: Marca de tiempo Unix de la ultima vez que se ejecutó la aplicación
- `language`: Idioma de la reseña.
- `review`: Contenido de la reseña.
- `timestamp_created`: Marca de tiempo Unix de la creación de la reseña.
- `timestamp_updated`: Marca de tiempo Unix cuando se actualizó por ultima vez.
- `voted_up`: Verdadero para una reseña positiva, falso en caso contrario.
- `votes_up`: Numero de usuarios que designaron la reseña como útil.
- `votes_funny`: Numero de usuarios que designaron la reseña como divertida.
- `weighted_vote_score`: Puntaje de utilidad de la reseña.
- `comment_count`: Cantidad de comentarios en respuesta a la reseña.
- `steam_purchase`: Verdadero en caso de que el usuario haya obtenido la aplicación mediante *Steam*
- `received_for_free`: Verdadero si el usuario designó el producto como obtenido de forma gratuita.
- `written_during_early_access`: Verdadero si la reseña fue realizada cuando el juego se encontraba en *Early Access*
- `developer_response`: Respuesta del desarrollador en caso de existir.
- `timestamp_dev_responded`: Marca de tiempo Unix al momento de responder el desarrollador.

Los datos retornados presentan una diversidad de información que sale del alcance abarcado, siendo los principales a utilizar el contenido de la reseña, la polaridad de esta, la marca de tiempo de actualización, y el cursor con su fundamental uso para la obtención de las reseñas.

3.3. Iteraciones

Para una explicación detallada y organizada de la generación del concepto de propuesta, se procede a agrupar los distintos pasos realizados en iteraciones con la respectiva idea principal a progresar.

3.3.1. Conocimiento de los datos

Como punto de partida, se busca un primer acercamiento al entendimiento de los datos mediante la utilización de métodos no supervisados de *machine learning* para una reseña respecto a su valoración.

Mediante la API de *Steam* (Valve Corporation, s.f.) se han obtenido 82965 reseñas correspondientes al idioma inglés del videojuego *Monster Hunter World*, para una manipulación sencilla y acorde solo se conservó el texto de la reseña y su recomendación positiva o negativa, la cual se puede comprender como la intención del sentimiento del usuario.

La limpieza de los datos consta de un reemplazo de las reseñas vacías por cadenas de texto vacías para evitar incompatibilidades por el tipo de variable, luego se tokeniza el texto para eliminar conectores, y se eliminan las reseñas con menos de 10 tokens.

Se utiliza Word2Vec como técnica de procesamiento de los datos, la cual se basa en *word embedding*, una técnica en el procesamiento del lenguaje donde las palabras son vectorizadas generalmente mediante el uso de redes neuronales, creando así una representación numérica de la relación entre las palabras. Los datos procesados se utilizan tanto para entrenamiento como de prueba y adicionalmente se seleccionan conceptos positivos y negativos generales incorporados en el vocabulario generado. Se crean dos modelos basados en los valores semánticos de los sentimientos referentes a los conceptos anteriormente mencionados, uno por el total y otro por los 30 conceptos más cercanos a estos.

Presentando exactitudes de 84.7590 % y 82.1546 % respectivamente los modelos constan de un buen indicio de certeza del sentimiento real indicado por el usuario. Adicionalmente se identifica que pese a que los datos fueron solicitados en inglés para mantener constancia y están etiquetados de esa manera al ser solicitados, se encuentran presente varias reseñas en otros idiomas incluyendo, por ejemplo, chino, el cual consta con su propio sistema de escritura provocando problemas en la codificación de los caracteres.

3.3.2. Modelo de grano fino

Como se ha mencionado anteriormente existe el *Fine-Grained Sentiment Analysis*, el cual trabaja la identificación del sentimiento de una reseña no solo como documento, sino a niveles de oración o característica, siendo esta última de vital importancia a la hora de indagar en el porque algunos videojuegos resultan más atractivos que otros. A consideración de lo anterior, la necesidad de ocupar un modelo de grano fino aplicable al caso se vuelve fundamental.

Se reutilizan las reseñas del videojuego *Monster Hunter World* en primera instancia ya que se prioriza la búsqueda de un modelo compatible a los datos ya conocidos.

La limpieza de los datos nuevamente consta de la conversión de una reseña vacía a cadena de texto, seguido en este caso de eliminar las *stop words* o palabras vacías.

PyABSA es un *framework* modularizado para facilitar la utilización de distintos modelos pre-entrenados de ABSA y sus sub-tareas (Yang et al., 2023), su instancia de uso consta de un modelo pre-entrenado aplicable a un *dataset* para afinarlo o utilizar un *checkpoint* para realizar el proceso. Se procede a extraer aspectos y los sentimientos respectivos utilizando el *checkpoint* 'english' del modelo FAST LCF-ATEPC.

Como resultado se obtienen 6227 aspectos, los cuales a su vez mencionan la cantidad de veces que fue identificado como positivo, negativo, u neutral. Se observa que los aspectos no son necesariamente únicos ya sea por errores de escritura, agrupación de palabras, o incluso la fuente del texto.

La siguiente iteración repite el procedimiento con la adición de la creación de un *dataset*, cuyo propósito es aplicarlo al modelo pre-entrenado logrando así un modelo de grano fino respecto al tema. Se transforman 1000 reseñas del videojuego para observar la compatibilidad e identificar su formato, para lo cual se utilizan los métodos proporcionados por PyABSA para la creación de los *datasets* del tipo APC y se transforman al tipo ATEPC. En contraste con el resultado anterior de 6227 aspectos en esta iteración se identifican 5991.

3.3.3. Afinación del modelo

En consideración del modelo de grano fino utilizable se observa la necesidad de afinar el modelo, mediante la recopilación de reseñas de juegos utilizados como referencia respecto a géneros en los cuales se encuentra el videojuego principal, del cual se extraerán aspectos de sus reseñas. Adicionalmente, se considera la búsqueda de una forma alternativa de crear el *dataset* APC.

Para esta nueva iteración se utilizan 1000 reseñas positivas y negativas de 8 juegos referentes a 4 etiquetas de géneros de videojuegos: Acción, Gran Banda Sonora, RPG, y Estrategia. Dando así 16000 reseñas en total para el *dataset* de afinación, mientras que el videojuego principal es *Library of Ruina* del cual se extraen 4941 reseñas siendo de estas 4727 positivas y 214 negativas.

Además de limpiar los datos mediante la eliminación de las *stops words* y las cadenas de texto vacías, se filtran reseñas ajenas al alfabeto latino mediante el uso de la codificación del formato UTF-8 para su decodificación en formato ASCII.

La creación del *dataset* APC utiliza la librería *Spacy* para aplicar procesamiento del lenguaje natural a las reseñas y separar estas en oraciones, a las cuales se les extrae los aspectos mediante la librería *Yake*, la cual cuenta con un método de extracción no supervisado, adaptable a cualquier texto u documento entregado, independiente del lenguaje humano, y capaz de incluir *stopwords* en la extracción de aspectos, entre otras características descritas en (Cam-

pos et al., 2018a),(Campos et al., 2018b), y (Campos et al., 2020). Se utiliza una razón 9:1 para entrenamiento y prueba respectivamente.

Como es observable en la figura 5 se logra identificar varias características del videojuego, incluso mencionando personajes pertenecientes a este como Angela o Roland.

	Negative	Neutral	Positive	Total
story	74	13	244	331
game	73	40	137	250
lobotomy	47	9	95	151
difficulty	103	9	29	141
fun	24	4	109	137
hours	52	10	39	101
library	43	11	43	97
music	14	1	77	92
good	10	1	69	80
hard	42	2	34	78
love	7	1	64	72
play	28	8	31	67
moon	11	6	46	63
roland	13	2	46	61
great	12	1	47	60
angela	26	7	23	56
worth	5	1	48	54
games	6	3	44	53
combat	9	11	29	49
early access	8	3	37	48

Figura 5: Top 20 Aspectos Identificados de *Library of Ruina*

3.3.4. Reestructuración

Con la identificación de diversos aspectos del videojuego, se idea agrupar estos en macro características como serían Jugabilidad, Audio, y Gráficos. Esto permitiría una comparación agrupada y simplificada entre estas características.

Las reseñas para el *dataset APC* están constituidas por 3 archivos de géneros de videojuegos, cada uno con el top 10 de ventas en *Steam* utilizando 1000 reseñas positivas e igual cantidad negativa. Mientras que el *dataset* a analizar esta vez consta de 23 juegos perteneciente al top

25 de ventas de videojuegos que cumplen con estar etiquetados como pertenecientes a los 3 géneros, siendo estos: Aventura, Fantasía, y RPG. La fecha de referencia para la información fue el 9 de Octubre del 2023.

La limpieza de los datos no difiere de la iteración anterior.

Se utiliza CAT por (Tulkens & van Cranenburgh, 2020), el cual realiza un método de asignación automática para identificar la categoría a la que pertenece el aspecto. Su funcionamiento es mediante *word embedding* para identificar la relación con las palabras designadas, en este caso, las características anteriormente mencionadas.

Se desestima su posible utilización debido a la repetición de términos entre características, y la similitud entre algunos términos como se observa en la figura 6, donde por ejemplo, el término *game* aparecido en las características *gameplay*, *graphics*, y *none*, con este ultimo contando con otro término similar, *great game*.

	Negative	Neutral	Positive	Total
("Baldur's Gate 3", 'none', 'game')	26	8	33	67
("Baldur's Gate 3", 'none', 'play')	17	7	20	44
("Baldur's Gate 3", 'none', 'fun')	7	0	18	25
("Baldur's Gate 3", 'none', 'hours')	12	4	8	24
("Baldur's Gate 3", 'graphics', 'game')	20	1	3	24
("Baldur's Gate 3", 'none', 'good')	3	1	20	24
("Baldur's Gate 3", 'none', 'played')	2	0	15	17
("Baldur's Gate 3", 'none', 'buy')	9	3	5	17
("Baldur's Gate 3", 'none', 'act')	6	6	4	16
("Baldur's Gate 3", 'none', 'fix')	13	1	0	14
("Baldur's Gate 3", 'none', 'rpg')	1	0	13	14
("Baldur's Gate 3", 'graphics', 'act')	9	1	2	12
("Baldur's Gate 3", 'graphics', 'fun')	7	0	5	12
("Baldur's Gate 3", 'none', 'amazing')	0	0	11	11
("Baldur's Gate 3", 'none', 'worth')	0	0	11	11
("Baldur's Gate 3", 'none', 'bugs')	11	0	0	11
("Baldur's Gate 3", 'none', 'great game')	5	0	6	11
("Baldur's Gate 3", 'gameplay', 'game')	2	1	7	10
("Baldur's Gate 3", 'none', 'baldur gate')	3	4	3	10
("Baldur's Gate 3", 'none', 'work')	7	1	2	10
("Baldur's Gate II: Enhanced Edition", 'none', 'baldur gate')	4	16	14	34
("Baldur's Gate II: Enhanced Edition", 'none', 'play')	8	6	16	30
("Baldur's Gate II: Enhanced Edition", 'none', 'classic')	3	6	19	28

Figura 6: Top 20 Aspectos Identificados y Clasificados de *Baldur's Gate 3*

3.3.5. Aplicación en un periodo de tiempo

El manejo del tiempo de las reseñas cobra importancia al pensar en el porque estos videojuegos se presentan en el top de ventas del día anteriormente mencionado.

Las reseñas ocupadas son las mismas que en el procedimiento anterior con el contraste que las pertenecientes al top 25 fueron filtradas para dejar solo reseñas desde el día 1 de septiembre del 2023, esto gracias a la marca de tiempo Unix incluida en las reseñas. No se producen modificaciones adicionales.

3.3.6. Variaciones

Adicionalmente al caso anterior, se consideran otra variación en las situaciones influyentes en el análisis: eliminar los videojuegos presentes en la realización del *dataset* APC. Se utiliza una versión en la cual las palabras vacías no se ven filtradas, pero presenta resultados contra-productivos a la identificación de los aspectos debido al ruido, por lo que se termina descartando.

3.4. Algoritmo Resultante

Después de las diversas iteraciones realizadas se presentan los siguientes códigos resultantes de las 3 fases en las que se dividió el algoritmo con su pertinente explicación, junto a una explicación del corpus resultante utilizado para los entrenamientos.

3.4.1. Corpus

El corpus para el caso dado se define como los datos a usar para el entrenamiento, lo cual se da al momento de afinar el modelo. En el proceso de iteraciones el corpus en bruto consta de un CSV con cada reseña almacenando sus datos de texto, polaridad, fecha de creación, fecha de actualización, y el identificador del videojuego al que pertenece. Además, para los datos a los que se les aplica el modelo afinado, se almacenan de la misma forma en caso de posible uso y se requiera una coherencia en el formato. A continuación se observan los valores estadísticos de las polaridades para las distintas clasificaciones utilizadas.

Tabla 2: Reseñas pertenecientes al top 25

Videojuego	Totales	Positivas	Negativas
Baldur's Gate 3	2000	1000	1000
New World	2000	1000	1000
The Elder Scrolls® Online	2000	1000	1000
FINAL FANTASY XIV Online	2000	1000	1000
Black Desert	2000	1000	1000
The Witcher® 3: Wild Hunt	2000	1000	1000
Disgaea 7: Vows of the Virtueless	172	172	0
Last Epoch	2000	1000	1000
Hogwarts Legacy	2000	1000	1000
The Elder Scrolls V: Skyrim Special Edition	2000	1000	1000
Monster Hunter: World	2000	1000	1000
Tiny Tina's Wonderlands	2000	1000	1000
DRAGON BALL Z: KAKAROT	2000	1000	1000
Conan Exiles	2099	1000	1099
SCARLET NEXUS	1782	1000	782
SWORD ART ONLINE Last Recollection	49	49	0
The Elder Scrolls Online: Necrom	94	94	0
Tales of Arise	2000	1000	1000
Sea of Stars	1223	1000	223
Baldur's Gate II: Enhanced Edition	1436	1000	436
Solasta: Crown of the Magister	2000	1000	1000
WitchSpring R	244	244	0
God of War	2000	1000	1000

Tabla 3: Reseñas Top 10 Aventura

Videojuego	Totales	Positivas	Negativas
Baldur's Gate 3	2000	1000	1000
Cyberpunk 2077	2000	1000	1000
New World	2000	1000	1000
Rust	2000	1000	1000
Warhammer 40,000: Darktide	2000	1000	1000
Grand Theft Auto V	2000	1000	1000
Starfield	2000	1000	1000
Tom Clancy's Ghost Recon® Breakpoint	2000	1000	1000
Lies of P	1464	1000	464
Remnant II	2000	1000	1000

Tabla 4: Reseñas Top 10 RPG

Videojuego	Totales	Positivas	Negativas
Baldur's Gate 3	2000	1001	999
Cyberpunk 2077	2000	1000	1000
ELDEN RING	2000	1000	1000
New World	2099	1000	1099
The Elder Scrolls V: Skyrim Special Edition	2000	1000	1000
Starfield	2000	1000	1000
Lies of P	1464	1000	464
Black Desert	2000	1000	1000
The Elder Scrolls® Online	2000	1000	1000
SCUM	2000	1000	1000

Tabla 5: Reseñas Top 10 Fantasía

Videojuego	Totales	Positivas	Negativas
Baldur's Gate 3	2000	1000	1000
ELDEN RING	2000	1000	1000
New World	2000	1000	1000
The Elder Scrolls V: Skyrim Special Edition	2000	1000	1000
Black Desert	2000	1000	1000
The Elder Scrolls® Online	2000	1000	1000
FINAL FANTASY XIV Online	2000	1000	1000
Hogwarts Legacy	2000	1000	1000
Total War: WARHAMMER III	2000	1000	1000

3.4.2. Obtención de los datos

La obtención de los datos se basa en la utilización de la API (Valve Corporation, s.f.), para obtener las reseñas de diversos juegos en el código 1. Entregando una lista de los identificadores utilizados por *Steam* para cada videojuego ordenados en su respectiva clasificación, se procede a realizar varias llamadas a la API, sin embargo, las solicitudes se deben regularizar realizándose cada 1.5 segundos para evitar una prohibición temporal por exceso de llamadas.

Para cada juego se busca idealmente almacenar 2000 reseñas con una mitad positiva y otra negativa, de esta forma obteniendo una información más representativa. Entre los datos almacenados se encuentran el texto de la reseña, su polaridad, el identificador del videojuego, y sus fechas de publicación y actualización. Una vez obtenida toda la información solicitada

se almacena en formato CSV para su posterior uso.

3.4.3. Creación del *dataset* APC o corpus

En el código 2 lo primero es la unión de los tres *datasets*, los cuales son los 10 videojuegos en el top de ventas de cada categoría correspondiente, siendo estas: Aventura, Fantasía, y RPG. Luego se le procede a la limpieza y se aplican procesos *NLP* mediante *Spacy* para separar el texto de las reseñas en oraciones. A continuación, se prepara un extractor de cinco *keywords* o palabras claves, se inicializa el clasificador, y se dividen los datos en una razón 9:1 para entrenamiento y prueba respectivamente.

La creación de los archivos se basa en la iteración de las oraciones identificadas, las cuales se procesan identificando las palabras claves existentes para identificar la polaridad del sentimiento, luego la oración se almacena en el archivo destacando la ubicación de la *keyword* junto al sentimiento. Es importante mencionar que tanto las oraciones como la palabra clave son limpiados adicionalmente en el paso anterior eliminando las palabras vacías, y se eliminan los casos donde la oración o la *keyword* desaparecen.

3.4.4. Entrenamiento y Ejecución

El código 3 consta de variaciones, el extractor de aspectos presenta la opción de realizar un entrenamiento con el *dataset* creado anteriormente o utilizando un *checkpoint* creado por un entrenamiento realizado previamente. Al cargar los datos del top 25 las variaciones presentes son la filtración de reseñas publicadas antes del 1 de septiembre del 2023 y la eliminación de videojuegos presentes en las clasificaciones utilizadas para la creación del *dataset* APC.

El resto del código no presenta alternaciones, se realiza la limpieza de datos correspondientes y se obtiene la lista de videojuegos en el *dataset*, para cada juego se utilizan el texto de las reseñas y se le extrae los aspectos junto con predecir la polaridad del sentimiento, luego se observa si la oración no se vuelve vacía con la utilización del texto limpio y no se utiliza en caso positivo.

En un diccionario se almacena por cada juego los aspectos identificados y la cantidad de veces mencionados en total y por polaridad.

CAPÍTULO 4

VALIDACIÓN DE LA SOLUCIÓN

La creación de un algoritmo el cual busca obtener datos implícitos no podría estar completo sin una validación, la cual no solo se realiza como método de comprobación de una coherencia entre los datos obtenidos, sino que en el presente caso también sirve como una forma de completar el dicho algoritmo al encontrar una relación constante.

4.1. Aprendizaje no supervisado y sus validaciones

Al momento de mencionar *machine learning* y validaciones, en general, se puede realizar mediante la separación de los datos en partes de entrenamiento y pruebas, lo que implica una necesidad de datos etiquetados para la comprobación de los resultados, pero dicho caso no sucede para los métodos no supervisados los cuales tienden a carecer de esta información clasificada a no ser que esto último sea realizado manualmente. Una opción de validación para casos no supervisados sería el análisis de variables obtenidas y verificar si existe una correlación, la cual ayude a entender el resultado y como este se relaciona a los datos.

Al mencionar las correlaciones, o más bien los coeficientes de correlación los cuales serían los valores obtenidos al realizar el análisis, se presentan los métodos de obtención de estas las cuales serían pruebas de estadísticas y estas pueden o no ser paramétricas. Estas pruebas buscan inferir una relación entre los datos, al observar si se ajustan a algún tipo de distribución conocida y pueda asignarsele una escala de intervalos en caso de la estadística paramétrica, con el caso de los no paramétricos no existiendo una asunción de pertenencia a dichas distribuciones, pero con ambos buscando una aceptación o un rechazo de una hipótesis previamente planteada.

Uno de los casos de utilización de la estadística no paramétrica es para casos en los que existen un orden de clasificación dado y se busca observar si existe una correlación con otros valores, siendo este el caso que se presenta con el algoritmo realizado, por lo que se procede a utilizar dos pruebas no paramétricas para la validación: El coeficiente de correlación de Spearman, y el coeficiente de correlación de rango de Kendall's.

4.2. Coeficiente de Correlación de Spearman

Para validar el trabajo realizado se utiliza el coeficiente de correlación de Spearman, el cual consiste en la obtención del valor ρ mediante la formula:

$$\rho = \frac{6 \sum d_i^2}{n(n^2 + 1)}$$

El valor ρ se interpreta como un numero indicando el tipo de correlación entre el rango de las posiciones de los valores a los que se le ha aplicado la formula, variando entre 1 y -1 para correlaciones positivas y negativas respectivamente, además de tener un valor mas cercano a 0 en caso de no existir la correlación. Los otros valores en la formula son n como el numero de parejas de valores comparados, y d_i la diferencia entre la posición de la pareja comparada.

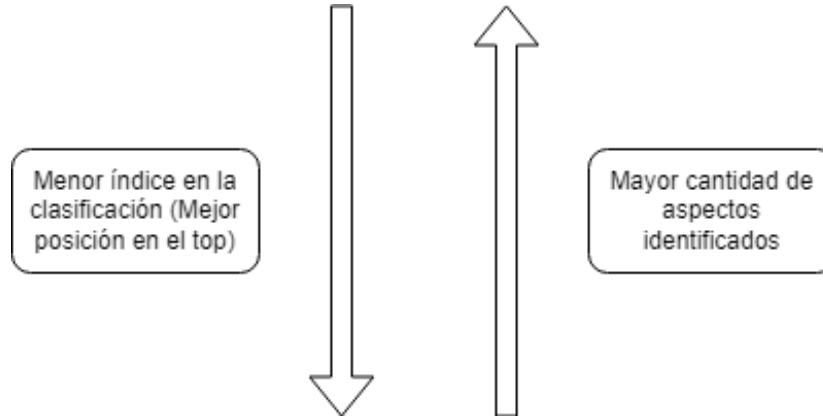


Figura 7: Explicación de la correlación negativa

4.3. Coeficiente de Correlación de Rango de Kendall

El llamado coeficiente de τ de Kendall, es otro valor calculable con el objetivo de identificar una correlación existente entre rangos de variables no paramétricas. Al igual que el coeficiente de Spearman su rango de valores oscila entre -1 y 1 y explica el tipo de relación identificada, siendo el caso del entero positivo una concordancia perfecta entre clasificaciones o en caso del entero negativo existiendo un desacuerdo o discordancia perfecta, lo cual se explica como una relación inversa. El valor 0 de la misma manera que en Spearman indicaría que las clasificaciones son independientes.

Para un par (X_i, Y_i) perteneciente a $(X_1, Y_1) \dots (X_n, Y_n)$, al compararse con un par (X_j, Y_j) tal que $i < j$, si la comparación tanto para el valor X como para el valor Y son menores o mayores se considera un par concordante, por lo que al ser distintos se considera discordante. Debido a que los empates no se consideran ni concordantes ni discordantes, la implementación de Tau a utilizar es τ_b cuya aplicación incluye ajustes a empates. La formula se describe como:

$$\tau_b = \frac{n_c - n_d}{\sqrt{(n_0 - n_1)(n_0 - n_2)}}$$

donde los valores n_c son la cantidad de pares concordantes, n_d son la cantidad de pares discordantes, y n_0, n_1, n_2 se definen como:

$$\begin{aligned}n_0 &= \frac{n(n-1)}{2} \\n_1 &= \sum_i \frac{t_i(t_i-1)}{2} \\n_2 &= \sum_j \frac{u_j(u_j-1)}{2}\end{aligned}$$

con t_i el numero de valores empatados en el i-ésimo grupo de empates para la primera cantidad y u_j el numero de valores empatados en el j-ésimo grupo de empates para la segunda cantidad.

Para su aplicación de los coeficientes de correlación en esta memoria se utilizan 2 listas, la primera correspondiendo a la posición de los videojuegos del top 25 y la segunda la cantidad de aspectos identificados de cada videojuego respectivamente. El conteo de aspectos se basa en la cantidad de veces identificado un aspecto junto al correspondiente sentimiento. En las siguientes tablas se observa el valor ρ y τ para el conteo de aspectos totales, positivos, neutrales, y negativos, junto a su respectivo valor P el cual es la probabilidad de rechazar la hipótesis nula la cual se considera como la instancia de no existir correlación. Como valores umbral para se usará $P \geq 0,1$ para rechazar la existencia de una relación, $0,1 > P \geq 0,01$ como una correlación no concluyente, y $0,01 > P$ para asumir la existencia de una correlación.

4.4. Caso Base

Con la ejecución del algoritmo en su estado base se obtiene la tabla 6, la cual resume las reseñas utilizadas por cada juego y la cantidad de aspectos identificados en esas reseñas con su respectiva polaridad.

Tabla 6: Aspectos Identificados

Videojuego	Reseñas	Aspectos			
		Totales	Positivos	Neutrales	Negativos
Baldur's Gate 3	1952	2931	904	327	1700
New World	1936	2948	759	375	1814
The Elder Scrolls® Online	1916	2777	871	390	1516
FINAL FANTASY XIV Online	1945	2787	729	372	1686
Black Desert	1911	2628	650	346	1632
The Witcher® 3: Wild Hunt	1873	2419	713	216	1490
Disgaea 7: Vows of the Virtueless	162	224	109	16	99
Last Epoch	1947	3001	915	304	1782
Hogwarts Legacy	1911	2744	841	380	1523
The Elder Scrolls V: Skyrim Special Edition	1930	2682	685	425	1572
Monster Hunter: World	1929	2370	640	282	1448
Tiny Tina's Wonderlands	1920	2713	922	340	1451
DRAGON BALL Z: KAKAROT	1927	2568	747	320	1501
Conan Exiles	2039	2929	815	253	1861
SCARLET NEXUS	1669	2243	855	225	1163
SWORD ART ONLINE Last Recollection	48	84	59	9	16
The Elder Scrolls Online: Necrom	88	98	56	19	23
Tales of Arise	1904	2572	1035	208	1329
Sea of Stars	1173	1663	1130	130	403
Baldur's Gate II: Enhanced Edition	1397	1914	848	276	790
Solasta: Crown of the Magister	1950	3121	1185	292	1644
WitchSpring R	223	359	249	39	71
God of War	1917	2403	776	255	1372

Tabla 7: Coeficiente de Correlación por Aspectos Identificados

Aspectos	Spearman		Kendall	
	ρ	Valor P	τ	Valor P
Totales	-0.474	0.0222	-0.359	0.0162
Positivos	0.0998	0.650	0.0513	0.754
Neutrales	-0.572	0.00433	-0.399	0.00720
Negativos	-0.577	0.00393	-0.415	0.00508

En la tabla 7 se observan valores ρ y τ , los cuales exceptuando por los aspectos positivos indican que a mejor posición en el top, se identifican una mayor cantidad de aspectos del respectivo sentimiento, resultando en una correlación negativa debido al funcionamiento del ranking en el cual un menor numero identifica una mayor posición, observable en la figura

7. A su vez, los valores P son extremadamente bajos y por ende se indica que efectivamente existiría una correlación en el caso Neutral y Negativo, mientras que en el caso Positivo queda descartado, además existiendo el caso Total en el cual dependiendo del valor de umbral α utilizado se podría rechazar o aceptar.

De forma alternativa se busca otro valor aplicable al coeficiente de correlación, surgiendo así la noción de una posible representación con menor influencia de parte de la cantidad de reseñas usadas finalmente por cada videojuego. En respuesta a la problemática, se opta por utilizar una razón entre la cantidad de aspectos identificados y la cantidad de reseñas, surgiendo la tabla 8. El uso del ratio finalmente explicaría si existe una correlación entre la clasificación de las ventas y la cantidad de aspectos de cada tipo identificados por reseña.

Tabla 8: Coeficiente de Correlación por ratio de Aspectos-Reseñas

Aspectos	Spearman		Kendall	
	ρ	Valor P	τ	Valor P
Totales	-0.146	0.505	-0.106	0.496
Positivos	0.541	0.00761	0.375	0.0118
Neutrales	-0.244	0.261	-0.169	0.270
Negativos	-0.646	0.000863	-0.470	0.00132

Entre los valores obtenidos de ρ y τ visibles en la tabla 8, se descartan los neutrales y totales ya que sus valores P son muy altos para ser tomados en consideración, por lo que se asume que no existe una correlación en esos ratios, mientras que respecto a los aspectos negativos al igual que en la tabla 7, presentan una correlación negativa indicando que a mejor posición en la clasificación por cada reseña se deberían identificar menos aspectos negativos. En contraste los aspectos positivos por reseña tienen una correlación positiva con la clasificación, presentando un caso contra intuitivo en el que a mayor índice clasificatorio se identifican mayor cantidad de aspectos positivos.

4.5. Caso Fecha Filtrada

Tabla 9: Coeficiente de Correlación por Aspectos Identificados (Fecha Filtrada)

Aspectos	Spearman		Kendall	
	ρ	Valor P	τ	Valor P
Totales	-0.516	0.0115	-0.367	0.0138
Positivos	-0.405	0.0547	-0.316	0.034
Neutrales	-0.578	0.00381	-0.403	0.00704
Negativos	-0.689	0.000271	-0.486	0.000866

Referente a la tabla 9 la cual considera reseñas del top 25 solo desde la fecha del 1 de septiembre del 2023 observa un aumento en los valores de correlación, siendo el aumento más

significativo el caso de los aspectos positivos. Los aumentos en los otros casos siguen la línea de los valores vistos en la tabla 7 siendo el caso de los neutrales el con menor diferencia a su valor anterior.

Tabla 10: Aspectos Identificados (desde 1 de Septiembre)

Videojuego	Reseñas	Aspectos			
		Totales	Positivos	Neutrales	Negativos
Baldur's Gate 3	1952	2931	904	327	1700
New World	1936	2948	759	375	1814
The Elder Scrolls® Online	702	912	387	133	392
FINAL FANTASY XIV Online	652	800	354	143	303
Black Desert	599	739	229	99	411
The Witcher® 3: Wild Hunt	1038	1009	569	113	327
Disgaea 7: Vows of the Virtueless	162	224	109	16	99
Last Epoch	934	1317	636	126	555
Hogwarts Legacy	1290	1653	726	268	659
The Elder Scrolls V: Skyrim Special Edition	1076	1240	533	253	454
Monster Hunter: World	1059	1066	512	174	380
Tiny Tina's Wonderlands	502	672	284	101	287
DRAGON BALL Z: KAKAROT	263	304	161	37	106
Conan Exiles	557	766	300	72	394
SCARLET NEXUS	76	109	59	8	42
SWORD ART ONLINE Last Recollection	48	84	59	9	16
The Elder Scrolls Online: Necrom	12	12	9	2	1
Tales of Arise	165	214	117	24	73
Sea of Stars	1146	1619	1127	126	366
Baldur's Gate II: Enhanced Edition	118	150	83	21	46
Solasta: Crown of the Magister	140	231	125	26	80
WitchSpring R	223	359	249	39	71
God of War	1008	947	591	118	238

Tabla 11: Coeficiente de Correlación por ratio de Aspectos-Reseñas (Fecha Filtrada)

Aspectos	Spearman		Kendall	
	ρ	Valor P	τ	Valor P
Totales	0.0464	0.833	0.027	0.875
Positivos	0.745	0.0000452	0.533	0.000217
Neutrales	-0.220	0.312	-0.154	0.319
Negativos	-0.593	0.00281	-0.462	0.00162

Respecto a la tabla 11, al igual que en tabla 8 la existencia de una correlación en los aspectos

totales y neutrales es rechazada, siendo como tal menos probable que en caso base. Así mismo para los aspectos negativos pese a tener un valor P más alto sigue siendo despreciable, por lo que sigue existiendo correlación, en cambio para los aspectos positivos se reafirma la existencia de una correlación debido a la disminución del valor P.

4.6. Caso Juegos Filtrados

Adicionalmente se realizan ejecuciones eliminando los videojuegos presentes en los archivos top 10 por géneros, esta edición solo elimina los videojuegos: *Baldur's Gate 3*, *New World*, *The Elder Scrolls® Online*, *FINAL FANTASY XIV Online*, *Black Desert*, *Hogwarts Legacy*, y *The Elder Scrolls V: Skyrim Special Edition*. No se produce ninguna alteración en las posiciones de la clasificación.

Tabla 12: Coeficiente de Correlación por Aspectos Identificados (Juegos Filtrados)

Aspectos	Spearman		Kendall	
	ρ	Valor P	τ	Valor P
Totales	-0.15	0.579	-0.0999	0.625
Positivos	0.217	0.418	0.166	0.398
Neutrales	-0.141	0.602	-0.0999	0.625
Negativos	-0.323	0.221	-0.183	0.350

Al analizar la tabla 12 no se logra observar ningún valor concluyente en alguno de los aspectos, asumiendo así la inexistencia de una correlación.

Tabla 13: Coeficiente de Correlación por ratio de Aspectos-Reseñas (Juegos Filtrados)

Aspectos	Spearman		Kendall	
	ρ	Valor P	τ	Valor P
Totales	0.179	0.506	0.166	0.398
Positivos	0.467	0.0677	0.366	0.0516
Neutrales	0.220	0.411	0.166	0.398
Negativos	-0.432	0.0944	-0.283	0.139

La tabla 13 presenta al igual que las tablas de ratios anteriores resultados no relacionados para aspectos totales y neutrales, mientras los aspectos positivos y negativos presentan valores de correlación que dan indicios de mantener la correlación en las razones de esos aspectos y la cantidad de reseñas, más sin ser concluyentes debido a los valores P y por lo tanto no se puede confirmar definitivamente la relación, siendo un caso similar para los valores τ con la diferencia que para los aspectos negativos superan el margen de rechazo para una posible correlación.

4.7. Caso Juegos y Fecha Filtrados

Finalmente se realiza un nuevo caso donde se ocupa tanto las fechas filtradas como los juegos filtrados, produciendo una situación cercana a una predicción de una clasificación.

Tabla 14: Coeficiente de Correlación por Aspectos Identificados (Juegos y Fecha Filtrados)

Aspectos	Spearman		Kendall	
	ρ	Valor P	τ	Valor P
Totales	-0.232	0.386	-0.15	0.450
Positivos	-0.133	0.620	-0.108	0.557
Neutrales	-0.166	0.538	-0.108	0.557
Negativos	-0.444	0.0848	-0.283	0.139

Al igual que en la tabla 12, la tabla 14 no presenta valores concluyentes para asumir correlación, pero en el caso de los aspectos negativos se observa un aumento en el valor de la correlación el cual no llega a ser irrefutable para ρ .

Tabla 15: Coeficiente de Correlación por ratio de Aspectos-Reseñas (Juegos y Fecha Filtrados)

Aspectos	Spearman		Kendall	
	ρ	Valor P	τ	Valor P
Totales	0.179	0.506	0.15	0.450
Positivos	0.561	0.0235	0.383	0.0412
Neutrales	0.332	0.208	0.216	0.265
Negativos	-0.399	0.124	-0.316	0.0960

Y la tabla 15 mantiene una estructura similar a tabla 13 con sus aspectos neutrales y positivos sin correlación, e indicios de tipos de correlación positiva y negativas para los aspectos del mismo nombre respectivamente. Se observa una mejoría al filtrar por fecha en los aspectos positivos, mientras que los negativos decaen levemente para ρ , y aumenta de forma leve para τ .

4.8. Correlación Aspectos Postivos-Negativos

Para finalizar se calcula un coeficiente de correlación adicional por cada caso revisado, este se realiza entre las posiciones de la clasificación del top 25 y el ratio entre aspectos positivos y negativos.

La tabla 16 resume los casos anteriores en donde al filtrar desde cierta fecha respecto al caso base la correlación se intensifica, mientras que al filtrar los juegos existentes en los datos del *dataset* APC pese a que esta relación se debilita se sigue manteniendo. Al utilizar el caso con los dos filtros pese a que la probabilidad de rechazarse tiene algo más de un valor del

Tabla 16: Coeficiente de Spearman por ratio de aspectos Positivos-Negativos

Caso	Spearman		Kendall	
	ρ	Valor P	τ	Valor P
Base	0.632	0.00120	0.438	0.00292
Fecha filtrada	0.779	0.0000115	0.588	0.0000340
Juegos filtrados	0.470	0.0658	0.3	0.116
Juegos y fecha filtrados	0.605	0.0128	0.399	0.0326

1 %, se puede mencionar como un valor significativo en lo que respecta a la utilización del algoritmo para lograr identificar aspectos, y clasificaciones respecto a estas. Adicionalmente se debe mencionar como los casos base y de fecha filtrada presentan valores concluyentes, con un valor no concluyente para ambos filtros aplicados, y el caso con solo los videojuegos filtrados siendo rechazado para Kendall a diferencia de Spearman, que presenta un valor inconcluyente. En la figura 8 se observa una explicación gráfica de la situación.

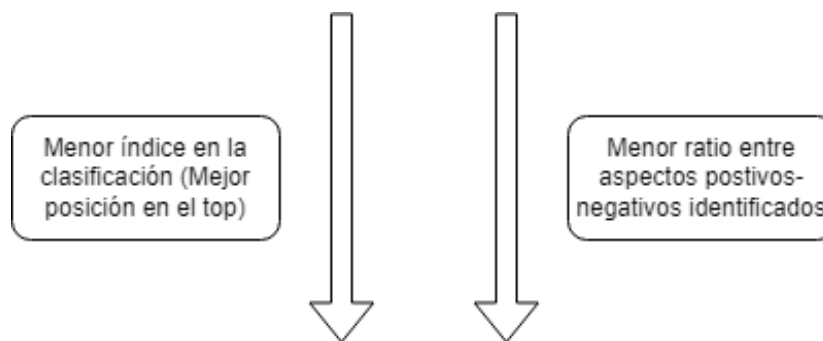


Figura 8: Correlación entre el Top de Ventas y el Ratio Positivo-Negativo

4.9. Comparación de coeficientes de correlación

Ambos coeficientes de correlación son no paramétricos, esto significa su utilización en casos donde no se puede prever una distribución. Al comparar los valores ρ y τ se observa que mantienen la misma positividad o negatividad respectiva. Mientras que respecto a los valores P se proceden variaciones mínimas en las probabilidades dándose pocos casos donde cambia la decisión de aceptación, incertidumbre, o rechazo de la hipótesis nula.

La tabla mas importante a mencionar es la tabla 16, la cual menciona el ratio entre los aspectos positivos y negativos, conteniendo valores significativos en los primeros 2 casos, mientras que los 2 restantes muestran resultados no concluyentes o al menos cercanos al umbral. Al observar ambos coeficientes de correlación obteniendo valores similares, aumenta la importancia otorgable al ratio investigado.

Debido a la naturaleza de los coeficientes, en caso de querer decantarse por uno esto se dificulta dandose la situación donde se vuelven un complementos del otro al reafirmar los resultados obtenidos.

CAPÍTULO 5

CONCLUSIONES

Se construyó un algoritmo no supervisado que permite la obtención de aspectos en una reseña, contando con la principal característica de no depender de datos etiquetados de referencia, lo cual facilita su reutilización con otros datos, pero complica su validación, y por ende, un índice de precisión.

Las técnicas utilizadas principalmente son la extracción de términos (ATE) y la clasificación de polaridad de aspectos (APC), las cuales son subtarefas del análisis de sentimientos basado en aspectos (ABSA), contando además con ejecuciones del procesamiento del lenguaje natural (NLP). Estas fueron seleccionadas en varias iteraciones al aplicar la metodología, debido a que el objetivo de encontrar una razón en el porqué del éxito fue asimilado a la búsqueda de características específicas, en contraste con el descartado *word embedding* que permite identificar las relaciones entre palabras.

Entre las limitaciones de su aplicación se encuentran: la necesidad de dos conjuntos de datos considerables con un manejo claro de las fechas de emisión de las reseñas, los largos periodos de entrenamiento y ejecución, y la falta de una validación que permita la obtención de la exactitud y precisión del algoritmo.

Respecto a los objetivos de esta memoria, el objetivo general busca la identificación de algún factor de influencia para decir que un videojuego es exitoso frente a sus reseñas analizadas, las cuales varían con el tiempo, considerando esto nace el ya mencionado ratio de aspectos positivos-negativos, el cual trabaja con los aspectos extraídos mediante análisis de sentimientos en las reseñas.

Para los objetivos específicos, el primer punto es sobre permitir una estimación del éxito del videojuego basándose en otros, lo cual se realiza de forma similar en el cuarto caso de cálculos de correlaciones, por lo que utilizando las fechas y géneros correctos se puede generar un posible top para comparar su éxito predicho.

El segundo punto indaga en el porqué se dan las reseñas de los jugadores, lo cual se puede mencionar como los aspectos identificados, más no categorizados, lo cual ayudaría a una mayor comprensión, sin mencionar la falta de implementación a periodos de tiempos como lo es el tiempo entre actualizaciones.

Finalmente, el tercer punto se resuelve de forma inherente con la realización del escrito y la utilización de *CRISP-DM* la cual permite iterar sobre las implementaciones para adaptarlas constantemente a las necesidades que surgen en el desarrollo.

En la realización del escrito y gracias a los coeficientes de correlación, se logra identificar y evidenciar una relación entre las reseñas y la popularidad en ventas de los videojuegos en

la plataforma *Steam*. Mediante una identificación automatizada de los aspectos existentes en las reseñas con su respectiva asignación de polaridad de sentimiento, se logra establecer una correlación entre su posición en la clasificación del top y la razón entre la cantidad de aspectos positivos y negativos identificados para el juego.

La existencia de esa correlación implicaría una forma de deducir una posible clasificación de juegos al ejecutar el algoritmo, extraer los aspectos de diversos juegos, y luego asignarlos en el potencial top de ventas. Pero debido a la designación de los valores umbral no es concluyente, requiriendo mejoras para disminuir la probabilidad de un falso positivo a la existencia de la correlación.

Se considera los usos de tanto el coeficiente de correlación de Spearman como el de Kendall, este ultimo contando adicionalmente con una menor sensibilidad a valores atípicos y pares empatados. La utilización de los dos coeficientes para casos no paramétricos permite reafirmar la probabilidad de existencia de la relación.

Durante el proceso de validación se han notado otros valores los cuales merece la pena mencionar y explicar. Entre ellos indiscutiblemente se ha de mencionar la disparidad entre reseñas disponibles y/o utilizadas para cada juego. Las tablas 6 y 10 muestran como al utilizar el filtro para ocupar reseñas desde la fecha especificada en promedio hubo una reducción del 50% de datos utilizables. Lo cual afectó positivamente la obtención de los valores del coeficiente de correlación aún pese a la disminución de datos.

Lo anterior da paso a dos interpretaciones, en primera instancia el uso de datos mas cercanos a la fecha de extracción de la clasificación deberían coincidir más apropiadamente con los motivos por los cuales se ha dado forma a ese orden, por otro lado, mientras este trabajo no considera un manejo de reseñas respecto a un periodo controlado de tiempo para la creación y utilización del *dataset* APC, el segmentarlo correctamente debería producir una mejora en los motivos del porque se ha formado tal top de ventas.

Otros valores importantes de mencionar son los coeficiente de correlación obtenidos por los ratios de aspectos y reseñas para los positivos y negativos, ya que en cada caso explorado los coeficientes siguen la misma tendencia a representar su polaridad, en el caso de los aspectos positivos, la correlación positiva indica que a mayor índice de posición en la clasificación se encuentran mas aspectos positivos por reseñas, siendo el caso contrario para los aspectos negativos los cuales van disminuyendo por reseña. Pese a lo contradictorio que puede ser que a los primeros lugares se les remarque más lo negativo que lo positivo, de ser así siguen siendo datos valiosos que les indican a los desarrolladores que aspectos deben mejorar y los cuales tienden a aparecer más a una mejor posición en el top de ventas.

En la realización del proceso se han identificado varias dificultades y/o alternativas, las cuales podrían influir y mejorar considerablemente el algoritmo además de la investigación realizada.

- Pese a que el tipo de practica realizada implica un uso contextual de los datos, la de-

tección de sarcasmo podría ser un elemento vital al momento de la utilización del sentimiento de la reseña al reasignar su polaridad, realizando así una lectura mas fiel a la interpretación de los datos.

- El volver a abarcar las macro categorías al identificar los aspectos podría permitir una comparación más pareja entre los videojuegos, pero el método de asignación a esas categorías debe ser pulido para evitar una distribución pobre o sin sentido.
- La normalización de contenidos de medios sociales podría ser una solución ante los dilemas de estandarización de aspectos, la repetición de estos en las macro categorías, o el texto mal escrito.
- Respecto a una observación detallada a un solo videojuego con tal de observar la variación en su éxito, con las modificaciones pertinentes a la segmentación de las reseñas en periodos de tiempos tales como actualizaciones, se puede obtener la correlación de los periodos mas productivos. A sí mismo, la obtención de los aspectos indicaría cuales fueron claves para el prestigio de la etapa, lo cual al combinarse con el punto anterior podría lograr una vista óptima de la evolución en la popularidad del juego.
- Pese a la identificación exitosa de una correlación, al considerar su dependencia de reseñas para la creación del *dataset* APC, se presenta la alta posibilidad de un periodo de validez para ese set de datos tras el cual deje de predecir apropiadamente, por lo que se instiga a la investigación de los posibles periodos de eficacia.
- Alternar y controlar la creación del *dataset* APC y el entrenamiento donde se utiliza, ya que se ha de mencionar que los tiempos de ejecución mediante *PyABSA* no son lineales, con la utilización de aproximadamente 60000 reseñas tardó 15 horas y 2 días respectivamente. Lo cual en contraste con la misma ejecución ocupando un 10 % de los datos tardaba aproximadamente 1 hora para las dos ejecuciones seguidas. A modo de referencia se ha utilizado una tarjeta gráfica o *GPU Nvidia* modelo RTX 3060ti 8Gb para el uso de núcleos *CUDA*, los cuales ayudan enormemente a la velocidad de procesamiento.
- Utilización de un modelo multilenguaje para permitir un abarque de una mayor cantidad de reseñas, y por lo tanto, de información disponible pero no utilizada.
- La posible utilización de distintas clasificaciones analizadas además puede proporcionar una cantidad de datos estudiabiles para observar si siguen algún tipo de distribución y por ende permitir el uso de la estadística paramétrica.
- La *API* utilizada para la obtención de los datos incluye variables no utilizadas, y estas podrían ayudar a la priorización de texto más substancialmente informativo mediante la utilización del puntaje de utilidad, o un filtrado de usuarios con un mínimo de tiempo en el videojuego.

ANEXOS

Código 1: Obtención de los Datos

```
import csv
import requests
import json
import time

def ApiCall(api, params=None):
    if params:
        response = requests.get(f"{api}", params=params)
    else:
        response = requests.get(f"{api}")
    return json.dumps(response.json(), sort_keys=True, indent=4)

name="data/top10-fantasy.csv"
games=[1086940,1245620,1063730,489830,582660,
        306130,39210,990080,1676840,1142710]
reviews = []
for game in games:

    cursor_last=None
    params={"num_per_page": "100",
            "filter": "updated",
            "review_type": "positive",
            "purchase_type": "all"
            }

    j=0
    alfa=True
    while True:
        time.sleep(1.5)
        json_ret=ApiCall("http://store.steampowered.com/appreviews/{id}?json=1".format(id=game))
        resp=json.loads(json_ret)
        j+=resp['query_summary']['num_reviews']
        for i in resp["reviews"]:
            reviews.append({"review": str(i["review"]), "sentiment": 1 if i["v"]>0 else -1})
        if resp["cursor"] == cursor_last:
            print(j)
            break
        else:
            print(j, game)
```

```
        cursor_last=resp["cursor"]
        params["cursor"]=cursor_last
        if j >=1000 and alfa:
            params["review_type"]="negative"
            params["cursor"]="*"
            alfa=False
        if j >=2000:
            break

with open(name, 'w', newline='', encoding="utf-8") as csvfile:
    fieldnames = ['review', 'sentiment', "game", "created", "updated"]
    writer = csv.DictWriter(csvfile, fieldnames=fieldnames)

    writer.writeheader()
    for i in reviews:
        writer.writerow(i)
```

Código 2: Creación del *Dataset*

```
import beepy
import pandas as pd
import os
import re
from nltk.corpus import stopwords
import warnings
from pyabsa import AspectPolarityClassification as APC, available_checkpo
import shutil
import spacy
import yake
from pyabsa.utils.absa_utils.absa_utils import convert_apc_set_to_atepc_set
warnings.filterwarnings("ignore")

data1 = pd.read_csv("data/top10-adventure.csv")
data2 = pd.read_csv("data/top10-fantasy.csv")
data3 = pd.read_csv("data/top10-rpg.csv")
data = pd.concat([data1, data2, data3], keys=["adventure", "fantasy", "rpg"])
data["review"] = data["review"].fillna(value="")

STOPWORDS = stopwords.words("english")
def clean(x):
    x = x.lower()
    x = re.sub("[^\w\d]", " ", x)
    x = " ".join(t for t in x.split() if t not in STOPWORDS)
```

```
        return x

def isEnglish(s):
    try:
        s.encode(encoding='utf-8').decode('ascii')
    except UnicodeDecodeError:
        return False
    else:
        return True

data["lang_en"]=data['review'].apply(lambda x : isEnglish(x))
data = data.drop(data[data['lang_en'] == False].index)

nlp = spacy.load("en_core_web_sm")
data["sent"]=data["review"].apply(nlp)

kw_extractor = yake.KeywordExtractor()
language = "en"
max_ngram_size = 2
deduplication_threshold = 0.1 #0.2
numOfKeywords = 5 #10
custom_kw_extractor = yake.KeywordExtractor(lan=language , n=max_ngram_size)

data_test = data.sample(frac = 0.1)
data_train = data.drop(data_test.index)

with open("E://Memoria/data/top20full.train.apc","w",encoding='UTF-8') as f:
    for doc in data_train["sent"]:
        for sent in doc.sents:
            line_data=clean(sent.text)
            keywords = custom_kw_extractor.extract_keywords(sent.text)
            if line_data not in (""," ")
and len(keywords)>0:
                for kw,_ in keywords:
                    kw_clean=clean(kw)
                    if kw_clean:
                        x=classifier.predict([line_data.replace(kw_clean, ' ')])
                        save_result=False,
                        print_result=False,
                        ignore_error=True,
                    )
                    fout.write(line_data.replace(kw_clean, "$T$")+"\n")
                    fout.write(kw_clean+"\n")
                    fout.write(x[0]["sentiment"][0]+" \n")
```

```
with open("E://Memoria/data/top20full.test.apc","w",encoding='UTF-8') as f:
    for doc in data_test["sent"]:
        for sent in doc.sents:
            line_data=clean(sent.text)
            keywords = custom_kw_extractor.extract_keywords(sent.text)
            if line_data not in (""," ")
and len(keywords)>0:
        for kw,_ in keywords:
            kw_clean=clean(kw)
            if kw_clean:
                x=classifier.predict([line_data.replace(kw_clean,'
                    save_result=False,
                    print_result=False,
                    ignore_error=True,
                ])
                fout.write(line_data.replace(kw_clean,"$T$")+"\n")
                fout.write(kw_clean+"\n")
                fout.write(x[0]["sentiment"][0]+"\\n")

convert_apc_set_to_atepc_set("E://Memoria/data/top20full.train.apc")
convert_apc_set_to_atepc_set("E://Memoria/data/top20full.test.apc")

src_file="E://Memoria/data/top20full.train.apc.atepc"
src_file_test="E://Memoria/data/top20full.train.apc.atepc"
dst_train="E://Memoria/integrated_datasets/atepc_datasets/21.top20full/top2
dst_test="E://Memoria/integrated_datasets/atepc_datasets/21.top20full/top2

if not os.path.exists("E://Memoria/integrated_datasets/atepc_datasets/21.top20full
    os.makedirs("E://Memoria/integrated_datasets/atepc_datasets/21.top20full

if os.path.exists(dst_train):
    os.remove(dst_train)
if os.path.exists(dst_test):
    os.remove(dst_test)

shutil.copy2(src_file,dst_train)
shutil.copy2(src_file_test,dst_test)
```

Código 3: Algoritmo

```
import numpy as np

def incmatrix(genl1,genl2):
import beepy
```

```
import pandas as pd # data processing , CSV file I/O (e.g. pd.read_csv)
import requests
from pyabsa import ATEPCCheckpointManager
import re
from nltk.corpus import stopwords
from pyabsa import ModelSaveOption , DeviceTypeOption
import warnings
from pyabsa import AspectTermExtraction as ATEPC
from scipy.stats import spearmanr
from scipy.stats import kendalltau
warnings.filterwarnings("ignore")

STOPWORDS = stopwords.words("english")
def clean(x):
    x = x.lower()
    x = re.sub("[^\w\d]", "_", x)
    x = "_".join(t for t in x.split() if t not in STOPWORDS)
    return x

def isEnglish(s):
    try:
        s.encode(encoding='utf-8').decode('ascii')
    except UnicodeDecodeError:
        return False
    else:
        return True

#aspect_extractor = ATEPCCheckpointManager.get_aspect_extractor(checkpoint
#                                                                auto_device=True
#                                                                )

config = (
    ATEPC.ATEPCConfigManager.get_atepc_config_english()
)
config.model = ATEPC.ATEPCModelList.FAST_LCF_ATEPC
config.batch_size = 16
config.patience = 2
config.log_step = -1
config.seed = [1]
config.verbose = False
config.notice = (
    "This is an training example for aspect term extraction"
)
trainer = ATEPC.ATEPCTrainer(
```

```
        config=config ,
        dataset="integrated_datasets/atepc_datasets/21.top20full" ,
        auto_device=DeviceTypeOption.AUTO,
        checkpoint_save_mode=ModelSaveOption.SAVE_MODEL_STATE_DICT ,
        load_aug=False ,
    )
    aspect_extractor = trainer.load_trained_model()

#data_fantasy = pd.read_csv("data/top10-fantasy.csv")
#data_adventure = pd.read_csv("data/top10-adventure.csv")
#data_rpg = pd.read_csv("data/top10-rpg.csv")
#data = pd.concat([data_adventure, data_fantasy, data_rpg], keys=["adventure", "fantasy", "rpg"])
#banlist=data["game"].unique()
data2 = pd.read_csv("data/top30-9_10-fantasy-rpg-adventure.csv")
#data2 = data2.drop(data2[data2['game'].isin(banlist)].index)
data2["review"] = data2["review"].fillna(value="")
data2["lang_en"]=data2['review'].apply(lambda x : isEnglish(x))
data2 = data2.drop(data2[data2['lang_en']==False].index)
#data2=data2.drop(data2[data2["updated"]<=1693540800].index)#1 sept 2023
data2['review_clean'] = data2['review'].apply(lambda x : clean(x))

def ApiCall(api, params=None):
    if params:
        response = requests.get(f"{api}", params=params)
    else:
        response = requests.get(f"{api}")
    return response.json()

games_list=data2["game"].unique()

games_names=[]
results_v2=dict()
for gm in games_list:
    api=ApiCall("https://store.steampowered.com/api/appdetails?appids={}".format(gm))
    games_names.append(api[str(gm)]["data"]["name"])
    text=list(data2["review_clean"][data2["game"]==gm])
    atepc_result = aspect_extractor.extract_aspect(inference_source=text,
                                                    pred_sentiment=True)

    sents=[]
    asp_sents=[]
    sents_sentiment=[]
    for i in range(len(atepc_result)):
        x=atepc_result[i]
```

```
        if x["sentence"].split():
            sents.append(x["sentence"].split())
            asp_sents.append(x["aspect"])
            sents_sentiment.append(x["sentiment"])
    results_v2[str(gm)] = dict()
    for j in range(len(sents)):
        for i in range(len(asp_sents[j])):
            if asp_sents[j][i] not in results_v2[str(gm)]:
                results_v2[str(gm)][asp_sents[j][i]] = {"Positive": 0, "Neutral": 0, "Negative": 0}
            results_v2[str(gm)][asp_sents[j][i]][sents_sentiment[j][i]] += 1
            results_v2[str(gm)][asp_sents[j][i]]["Total"] += 1

lista = []
lista_pos = []
lista_neg = []
lista_neu = []
lista_total = []
ratio_total = []
ratio_pos = []
ratio_neu = []
ratio_neg = []
ratio_p_n = []
for i in range(len(games_list)):
    revs = data2[data2["game"] == games_list[i]].shape[0]
    pos = 0
    neg = 0
    neu = 0
    total = 0
    for x in results_v2[str(games_list[i])]:
        pos += results_v2[str(games_list[i])][x]["Positive"]
        neg += results_v2[str(games_list[i])][x]["Negative"]
        neu += results_v2[str(games_list[i])][x]["Neutral"]
        total += results_v2[str(games_list[i])][x]["Total"]
    lista.append(i)
    lista_pos.append(pos)
    lista_neg.append(neg)
    lista_neu.append(neu)
    lista_total.append(total)
    ratio_total.append(total/revs)
    ratio_pos.append(pos/revs)
    ratio_neu.append(neu/revs)
    ratio_neg.append(neg/revs)
    ratio_p_n.append(pos/neg)
```

```
print("Spearman")

print("Aspectos")
print("Total",spearmanr(lista,lista_total))
print("Pos",spearmanr(lista,lista_pos))
print("Neu",spearmanr(lista,lista_neu))
print("Neg",spearmanr(lista,lista_neg))

print("Ratios")
print("Pos",spearmanr(lista, ratio_pos))
print("Neg",spearmanr(lista, ratio_neg))
print("Neu",spearmanr(lista, ratio_neu))
print("Total",spearmanr(lista, ratio_total))

print("Ratio_Pos:Neg")
print("Pos:Neg",spearmanr(lista, ratio_p_n))

print("Kendall")

print("Aspectos")
print("Total",kendalltau(lista,lista_total))
print("Pos",kendalltau(lista,lista_pos))
print("Neu",kendalltau(lista,lista_neu))
print("Neg",kendalltau(lista,lista_neg))

print("Ratios")
print("Total",kendalltau(lista, ratio_total))
print("Pos",kendalltau(lista, ratio_pos))
print("Neu",kendalltau(lista, ratio_neu))
print("Neg",kendalltau(lista, ratio_neg))

print("Ratio_Pos:Neg")
print("Pos:Neg",kendalltau(lista, ratio_p_n))
```

REFERENCIAS BIBLIOGRÁFICAS

- Andreev, G., Saxena, D., & Verma, J. K. (2021). Impact of Review Sentiment and Magnitude on Customers' Recommendations for Video Games. *2021 International Conference on Computational Performance Evaluation (ComPE)*, 992-995. <https://doi.org/10.1109/ComPE53109.2021.9752380>
- ARIK, K. (2022). Social Media Content Review of Popular MMORPG Games: Reddit Comment Scraping and Sentiment Analysis. *Journal of Emerging Computer Technologies*, 2(1), 13-21.
- Bhakuni, M., Kumar, K., Sonia, Iwendi, C., & Singh, A. (2022). Evolution and Evaluation: Sarcasm Analysis for Twitter Data Using Sentiment Analysis (S. Bhattacharya, Ed.). *Journal of Sensors*, 2022, 1-10. <https://doi.org/10.1155/2022/6287559>
- Birkett, J. (2018, 4 de mayo). Using Steam reviews to estimate sales. Consultado el 1 de noviembre de 2022, desde <https://www.gamedeveloper.com/business/using-steam-reviews-to-estimate-sales>
- Campos, R., Mangaravite, V., Pasquali, A., Jorge, A., Nunes, C., & Jatowt, A. (2020). YAKE! Keyword extraction from single documents using multiple local features. *Information Sciences*, 509, 257-289. <https://doi.org/10.1016/j.ins.2019.09.013>
- Campos, R., Mangaravite, V., Pasquali, A., Jorge, A. M., Nunes, C., & Jatowt, A. (2018a). A Text Feature Based Automatic Keyword Extraction Method for Single Documents. En G. Pasi, B. Piwowarski, L. Azzopardi & A. Hanbury (Eds.), *Advances in Information Retrieval* (pp. 684-691). Springer International Publishing. https://doi.org/10.1007/978-3-319-76941-7_63
- Campos, R., Mangaravite, V., Pasquali, A., Jorge, A. M., Nunes, C., & Jatowt, A. (2018b). YAKE! Collection-Independent Automatic Keyword Extractor. En G. Pasi, B. Piwowarski, L. Azzopardi & A. Hanbury (Eds.), *Advances in Information Retrieval* (pp. 806-810). Springer International Publishing. https://doi.org/10.1007/978-3-319-76941-7_8
- Chakraborty, S., Mobin, I., Roy, A., & Khan, M. H. (2018). Rating Generation of Video Games using Sentiment Analysis and Contextual Polarity from Microblog. *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, 157-161. <https://doi.org/10.1109/CTEMS.2018.8769149>
- Chen, X., & Wan, X. (2022). A Simple Information-Based Approach to Unsupervised Domain-Adaptive Aspect-Based Sentiment Analysis. *CoRR*, abs/2201.12549. <https://arxiv.org/abs/2201.12549>
- Cortis, K., & Davis, B. (2022, junio). Baseline English and Maltese-English Classification Models for Subjectivity Detection, Sentiment Analysis, Emotion Analysis, Sarcasm Detection, and Irony Detection. En M. Melero, S. Sakti & C. Soria (Eds.), *Proceedings of the 1st Annual Meeting of the ELRA/ISCA Special Interest Group on Under-Resourced Languages* (pp. 161-168). European Language Resources Association. <https://aclanthology.org/2022.sigul-1.21>

- Dylan. (2019, noviembre). How to plan a Machine Learning project? Try CRISP-DM. Consultado el 8 de diciembre de 2022, desde <https://www.nimblecoding.com/data-mining-process-crisp-dm/>
- Eberhard, L., Kasper, P., Koncar, P., & Gutl, C. (2018). Investigating Helpfulness of Video Game Reviews on the Steam Platform. *2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, 43-50. <https://doi.org/10.1109/SNAMS.2018.8554542>
- Guo, X., Yu, W., & Wang, X. (2021). An Overview on Fine-grained Text Sentiment Analysis: Survey and Challenges. *Journal of Physics: Conference Series*, 1757, 012038. <https://doi.org/10.1088/1742-6596/1757/1/012038>
- Gupta, S., Singh, R., & Singla, V. (2020). *Emoticon and Text Sarcasm Detection in Sentiment Analysis* (J. Arpad, P. R. Chandra, G. Xiao-Zhi, S. D. L. A. Kumar & Kosa, Eds.). https://doi.org/10.1007/978-981-15-0029-9_1
- Guzsvinecz, T. (2023). The correlation between positive reviews, playtime, design and game mechanics in souls-like role-playing video games. *Multimedia Tools and Applications*, 82, 4641-4670. <https://doi.org/10.1007/s11042-022-12308-1>
- Kang, H.-N. (2017). A Study of Analyzing on Online Game Reviews Using a Data Mining Approach: STEAM Community Data. *International Journal of Innovation, Management and Technology*, 90-94. <https://doi.org/10.18178/ijimt.2017.8.2.709>
- Kasper, P., Koncar, P., Santos, T., & Gütl, C. (2019). On the Role of Score, Genre and Text in Helpfulness of Video Game Reviews on Metacritic. *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, 75-82. <https://doi.org/10.1109/SNAMS.2019.8931866>
- Kewirata, P., Sureerat, S., Buachai, W., Rujivorakul, V., Jakkaew, P., & Intayoad, W. (2022). Exploring Players' Sentiments Toward Simulation Games: A Case Study in Thailand. *2022 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*, 467-471. <https://doi.org/10.1109/ECTIDAMTNCON53731.2022.9720365>
- Kotelnikova, A. V., Vychezhzhanin, S. V., & Kotelnikov, E. V. (2023). Cross-Domain Sentiment Analysis Based on Small in-Domain Fine-Tuning. *IEEE Access*, 11, 41061-41074. <https://doi.org/10.1109/ACCESS.2023.3269720>
- Kusumawardani, R. P., Priansya, S., & Atletiko, F. J. (2018). Context-sensitive normalization of social media text in bahasa Indonesia based on neural word embeddings [INNS Conference on Big Data and Deep Learning]. *Procedia Computer Science*, 144, 105-117. <https://doi.org/https://doi.org/10.1016/j.procs.2018.10.510>
- Li, X., Wang, X., & Liu, H. (2021). Research on fine-tuning strategy of sentiment analysis model based on BERT. *2021 International Conference on Communications, Information System and Computer Engineering (CISCE)*, 798-802. <https://doi.org/10.1109/CISCE52179.2021.9445882>
- Lin, D., Bezemer, C.-P., Zou, Y., & Hassan, A. E. (2019). An empirical study of game reviews on the Steam platform. *Empirical Software Engineering*, 24, 170-207. <https://doi.org/10.1007/s10664-018-9627-4>

- Nandi, B. P., Jain, A., & Tayal, D. K. (2023). Aspect Based Sentiment Analysis Using Long-Short Term Memory and Weighted N-Gram Graph-Cut. *Cognitive Computation*, 15, 822-837. <https://doi.org/10.1007/s12559-022-10104-5>
- Nayel, H., Amer, E., Allam, A., & Abdallah, H. (2021, abril). Machine Learning-Based Model for Sentiment and Sarcasm Detection. En N. Habash, H. Bouamor, H. Hajj, W. Magdy, W. Zaghouni, F. Bougares, N. Tomeh, I. Abu Farha & S. Touileb (Eds.), *Proceedings of the Sixth Arabic Natural Language Processing Workshop* (pp. 386-389). Association for Computational Linguistics. <https://aclanthology.org/2021.wanlp-1.51>
- Oh, S., Lee, D., Whang, T., Park, I., Gaeun, S., Kim, E., & Kim, H. (2021). Deep Context- and Relation-Aware Learning for Aspect-based Sentiment Analysis. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, 495-503. <https://doi.org/10.18653/v1/2021.acl-short.63>
- Punetha, N., & Jain, G. (2023a). Aspect and orientation-based sentiment analysis of customer feedback using mathematical optimization models. *Knowledge and Information Systems*, 65, 2731-2760. <https://doi.org/10.1007/s10115-023-01848-z>
- Punetha, N., & Jain, G. (2023b). Bayesian game model based unsupervised sentiment analysis of product reviews. *Expert Systems with Applications*, 214, 119128. <https://doi.org/10.1016/j.eswa.2022.119128>
- Ruseti, S., Sirbu, M.-D., Calin, M. A., Dascalu, M., Trausan-Matu, S., & Militaru, G. (2020). Comprehensive Exploration of Game Reviews Extraction and Opinion Mining Using NLP Techniques. En X.-S. Yang, S. Sherratt, N. Dey & A. Joshi (Eds.), *Fourth International Congress on Information and Communication Technology* (pp. 323-331). Springer Singapore. https://doi.org/https://doi.org/10.1007/978-981-15-0637-6_27
- Santos, R., Arrais, J. P., & Silva, P. A. (2021). Analysing Games for Health through Users' Opinion Mining. *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, 319-323. <https://doi.org/10.1109/CBMS52027.2021.00035>
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science*, 181, 526-534. <https://doi.org/10.1016/j.procs.2021.01.199>
- Sehar, U., Kanwal, S., Dashtipur, K., Mir, U., Abbasi, U., & Khan, F. (2021). Urdu Sentiment Analysis via Multimodal Data Mining Based on Deep Learning Algorithms. *IEEE Access*, 9, 153072-153082. <https://doi.org/10.1109/ACCESS.2021.3122025>
- Shangipour ataei, T., Javdan, S., & Minaei-Bidgoli, B. (2020, julio). Applying Transformers and Aspect-based Sentiment Analysis approaches on Sarcasm Detection. En B. B. Klebanov, E. Shutova, P. Lichtenstein, S. Muresan, C. Wee, A. Feldman & D. Ghosh (Eds.), *Proceedings of the Second Workshop on Figurative Language Processing* (pp. 67-71). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.figlang-1.9>
- Sngular. (2016, agosto). CRISP-DM: La metodología para poner orden en los proyectos. Consultado el 8 de diciembre de 2022, desde <https://www.sngular.com/es/data-science-crisp-dm-metodologia/>
- Steam Store. (s.f.). Consultado el 4 de diciembre de 2023, desde <https://store.steampowered.com/>

- Strååt, B., & Verhagen, H. (2017). Using User Created Game Reviews for Sentiment Analysis: A Method for Researching User Attitudes. *GHITALY@ CHIItaly*.
- Sykownik, P., Born, F., & Masuch, M. (2019). Can You Hear the Player Experience? A Pipeline for Automated Sentiment Analysis of Player Speech. *2019 IEEE Conference on Games (CoG)*, 1-4. <https://doi.org/10.1109/CIG.2019.8848096>
- Touahri, I., & Mazroui, A. (2021). Enhancement of a multi-dialectal sentiment analysis system by the detection of the implied sarcastic features. *Knowledge-Based Systems*, 227, 107232. <https://doi.org/https://doi.org/10.1016/j.knosys.2021.107232>
- Tulkens, S., & van Cranenburgh, A. (2020, julio). Embarrassingly Simple Unsupervised Aspect Extraction. En D. Jurafsky, J. Chai, N. Schluter & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 3182-3187). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.290>
- Urriza, I. M., & Clarino, M. A. A. (2021). Aspect-Based Sentiment Analysis of User Created Game Reviews. *2021 24th Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA)*, 76-81. <https://doi.org/10.1109/O-COCOSDA202152914.2021.9660559>
- Valve Corporation. (s.f.). User Reviews - Get List (Steamworks Documentation). Consultado el 1 de noviembre de 2022, desde <https://partner.steamgames.com/doc/store/getreviews>
- Valve Corporation. (2017). Steam Direct Fee (Steamworks Documentation). Consultado el 23 de noviembre de 2023, desde <https://partner.steamgames.com/doc/gettingstarted/appfee>
- Vashishtha, S., & Susan, S. (2019). Fuzzy rule based unsupervised sentiment analysis from social media posts. *Expert Systems with Applications*, 138, 112834. <https://doi.org/https://doi.org/10.1016/j.eswa.2019.112834>
- Verma, P., Shukla, N., & Shukla, A. (2021). Techniques of Sarcasm Detection: A Review. *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 968-972. <https://doi.org/10.1109/ICACITE51222.2021.9404585>
- Video Game Insights. (2020, 9 de agosto). Infographic: Indie game revenues on Steam. Consultado el 1 de noviembre de 2022, desde <https://vginsights.com/insights/article/infographic-indie-game-revenues-on-steam>
- Viggiato, M., Lin, D., Hindle, A., & Bezemer, C.-P. (2022). What Causes Wrong Sentiment Classifications of Game Reviews? *IEEE Transactions on Games*, 14, 350-363. <https://doi.org/10.1109/TG.2021.3072545>
- Yang, H., Zeng, B., Yang, J., Song, Y., & Xu, R. (2021). A multi-task learning model for Chinese-oriented aspect polarity classification and aspect term extraction. *Neurocomputing*, 419, 344-356. <https://doi.org/https://doi.org/10.1016/j.neucom.2020.08.001>
- Yang, H., Zhang, C., & Li, K. (2023). PyABSA: A Modularized Framework for Reproducible Aspect-based Sentiment Analysis. En I. Frommholz, F. Hopfgartner, M. Lee, M. Oakes, M. Lalmas, M. Zhang & R. L. T. Santos (Eds.), *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, CIKM 2023, Bir-*

- mingham, United Kingdom, October 21-25, 2023* (pp. 5117-5122). ACM. <https://doi.org/10.1145/3583780.3614752>
- Zainuddin, N., Selamat, A., & Ibrahim, R. (2018). Hybrid sentiment classification on twitter aspect-based sentiment analysis. *Applied Intelligence*, 48, 1218-1232. <https://doi.org/10.1007/s10489-017-1098-6>