

Análisis de sentimientos en videojuegos

Daniel Iván Angeles Gómez¹
angodi228@gmail.com

Maricela Quintana López¹
mqintanal@uaemex.mx

¹Centro Universitario UAEM Valle de México
Universidad Autónoma del Estado de México

Resumen.

Los videojuegos multijugador han existido desde la década de los setenta, pero no es hasta mediados de los noventa, con la evolución de las redes que permiten la comunicación desde cualquier parte del mundo, que adquieren la característica de jugarse en línea, dándole otra dimensión humana al juego ya que se requiere una manera de comunicarse para formular estrategias o simplemente para platicar mientras se juega. La comunicación se puede realizar vía voz o texto a través de una interfaz, y a lo largo del juego se pueden tener conversaciones que van desde las casuales, que tienen que ver con aspectos generales del juego, hasta las agresivas, que incluyen insultos a los demás jugadores. En este artículo se muestran los resultados de aplicar el análisis de sentimientos a 75 conversaciones extraídas de partidas de los videojuegos League of Legends, Overwatch y Heroes of the Storm, 25 están en idioma español, y 50 en inglés; las conversaciones son de longitud variable, la más extensa contiene 440 palabras, mientras que la más pequeña contiene solo 50 palabras. Utilizando Semantría para Excel, un API del software Lexalytics, se creó un diccionario que pudiera ser utilizado para la clasificación de las conversaciones. Las clases se definieron como positivas, negativas o neutrales. Comparando el resultado con la clase dada por el experto, 47 de 75 conversaciones fueron clasificadas correctamente, dando un 62.6% de certeza a la clasificación, en búsqueda de aumentar la certeza de la clasificación se hizo uso de el algoritmo C4.5, en el cual los resultados mostraron el entorno de los implicados en las conversaciones, la certeza se considera podría mejorarse al incluir más conversaciones y un mejor balance en las clases.

Palabras clave: Análisis de sentimientos, clasificación, videojuegos.

Abstract

Multiplayer videogames exist since the 70's (seventies), but it isn't until mid 90's (nineties) with the evolution of networks that allow communication from anywhere in the world, that videogames become online games, giving them a new human dimension by requiring a way to communicate in order to create strategies, or just talk while playing. The communication can be via voice or text lines within an interface, and throughout the game, conversations can go from casual topics related to the general aspects of the game, to aggressive conversations insulting other players. This article shows the results of a Sentiments Analysis of 75 conversations extracted from the videogames: League of Legends, Overwatch and Heroes of the Storm; 25 in Spanish language, and 50 in English. The conversations are of variable length, the longest has 440 words, while the shortest only has 50 words. Using Semantría for Excel, an API software of Lexalytics, a lexicon dictionary was created in order to classify the conversations.

The classes were defined as positive, negative or neutral. Comparing these results with the given ones by an expert, 47 of the 75 conversations were correctly classified, getting a 62.6% classification accuracy, which can be improved using the C4.5 algorithm, the results showed the environment in the conversations players, the accuracy can be better with more conversations and a higher class balance.

Keywords: Sentiment Analysis, Classification, Videogames.

1 Introducción

Los videojuegos han tenido un creciente impacto en la industria del entretenimiento gracias a los dispositivos móviles, ya que estos permiten el uso de diferentes tipos de juegos, que van desde rompecabezas (“puzzles”), sin un fin específico (“arcade”), de estrategia (“RTC” Real Time Combat), hasta juegos donde su historia es inmersiva llegando a los niveles de un filme cinematográfico. Su gran diversidad ha permitido que personas de todas las edades los jueguen (Harris, 2014).

Hoy en día, los videojuegos, han llegado a formar parte fundamental en la rutina diaria de muchas personas, ya que están presentes como medios de entretenimiento y en algunos casos como vida laboral, esto incluye, a los que se dedican a desarrollar los videojuegos y a los jugadores profesionales que participan en diferentes eventos obteniendo una remuneración (Steven, 2016).

Si bien desde un principio, han existido juegos para uno o varios jugadores, no es hasta que la tecnología, particularmente las redes, evolucionaron al punto en el que la comunicación podía realizarse desde cualquier parte del mundo, que los juegos multijugador adquieren la característica de poderse jugar en línea (“online”), dándole otra dimensión humana al juego ya que se requiere una manera de comunicarse para formular estrategias o simplemente para platicar mientras se juega. La comunicación se puede realizar vía voz o texto a través de una interfaz, y a lo largo del juego se pueden tener conversaciones que van desde las casuales, que tienen que ver con aspectos generales del juego, hasta las agresivas, que incluyen insultos a los demás jugadores.

Dada la gran variedad de gente que interactúa en los videojuegos, existe una inmensa variedad de conversaciones que reflejan cómo se comportan los jugadores a lo largo de una partida en diferentes circunstancias del juego, desde cuándo van ganando hasta cuando la derrota es inminente, por lo que dichas conversaciones pueden analizarse utilizando la minería de opinión o análisis de sentimientos.

A diferencia de otras aplicaciones en donde todas las conversaciones tienden a ser de dominio público, en el caso particular de los videojuegos se carece del acceso directo a las conversaciones, por ende, es más complicado el extraer la información. En algunas ocasiones, los juegos varían en algunos aspectos dependiendo de la región, entre estos, el doblaje del juego, la ambientación y posibles temas sensibles, por ejemplo, el juego Wolfenstein tiene referencias al régimen nazi, por lo que en Alemania los símbolos o personajes contienen variaciones. Otro aspecto relacionado con la región es el servidor donde se alojan las partidas.

En este trabajo, se presentan los resultados del análisis de sentimientos realizado utilizando las conversaciones de tres juegos que tienen una versión similar para todo el mundo y solo varían en el doblaje: League of Legends (“LOL”), Overwatch (“OVW”) y Heroes of the Storm (“HOTS”), en el caso del primero se ocupará el servidor LAN (Latino América) y en los últimos dos el servidor de las “AMERICAS” que ocupa todo el continente americano.

El documento está organizado como sigue, en la sección 2 se presenta el análisis de sentimientos y la metodología que se utiliza. En la sección 3 se presentan las etapas realizadas para el análisis de sentimientos en videojuegos. Finalmente en la sección 4, se presentan las conclusiones y trabajo futuro.

2 Análisis de sentimientos

El análisis de sentimientos, SA, por sus siglas en inglés Sentiment Analysis, es el estudio de conversaciones donde se incluyen opiniones, actitudes y emociones realizado por computadora, puede ir enfocado a conversaciones individuales, hacia un tema o ciertos tipos de eventos, el SA se desempeña en diferentes niveles, pueden ser, a nivel de documento, oración o aspecto, en el primer nivel refleja los sentimientos expresados en un documento que son clasificados como positivos o negativos, en el segundo caso se realiza en base a oraciones que llegan a ser de poca longitud, existe mucha similitud en ambos casos debido a que las oraciones son subconjuntos de un documento, por último se tiene el análisis por aspecto, en los aspectos se intenta encontrar palabras claves que puedan ser de gran impacto en una oración o documento (Medhat, Hassan, & Korashy, 2014).

Los campos más explotados dentro del análisis de sentimientos están comúnmente relacionados con las redes sociales y con las reseñas de productos o servicios, los estudios presentan la opinión de las personas en base a un tema o dirigido a un producto, comúnmente clasificado en Positivo, Neutral o Negativo, los análisis que han tenido mayor mención se han desarrollado alrededor de Twitter, cada estudio tiene sus propias características como puede ser la metodología aplicada, las herramientas o la cantidad de datos.

Dentro del campo del análisis de sentimientos, se encuentra el SemEval-2016 Task 4: Sentiment Analysis in Twitter, un estudio que pretende averiguar los temas positivos o negativos que han surgido en Twitter. Una de las principales costumbres en este tipo de análisis es tener dos parámetros para los sentimientos: NEGATIVO y POSITIVO, sin embargo, SemEval-2016 Task 4, debido a la gran cantidad de datos e información que manejan, añade tres evaluaciones más, teniendo como parámetros finales: MUY POSITIVO, POSITIVO, NEUTRAL, NEGATIVO y MUY NEGATIVO, esto con el fin de que al tener más evaluaciones se tiene un mayor grado de exactitud al momento de clasificar las conversaciones, se utilizaron diferentes herramientas para comparar la certeza de cada una. El peor resultado que se registró fue de 37.4% de asertividad mientras que el más alto obtuvo una puntuación de 64.6% (Navok, Ritter, Rosenthal, Sebastiani, & Stoyanov, 2016).

Un año más tarde se realizó nuevamente el estudio SemEval-2017 Task 4, se ocuparon metodologías y procedimientos similares. Las principales modificaciones que se realizaron fueron las actualizaciones de algunas herramientas, se añadieron nuevos sistemas y hubo incorporación de diferentes lenguajes como el arábico o el español, dentro de la clasificación se observaron diferentes mejoras en la clasificación, donde la más baja obtuvo 20.6% de asertividad mientras que el más alto consiguió un total de 89.7% (Rosenthal, Farra, & Nakov, 2017).

Existen muchas formas de evaluar una conversación o frases dentro del SA, la más usual es evaluar toda la conversación sin modificar en ningún momento su estructura, sin embargo, también se pueden retirar cierta parte de la información como pre-procesamiento, tal como demostraron (Mat Zin, Mustapha, Azmi Murad, & Mohd Sharef, 2017) en su artículo donde analizan reseñas de películas. Parte de la metodología, fue retirar contenido para centrarse en las palabras o frases en las conversaciones, el contenido retirado fueron artículos, palabras sin sentido, fechas, caracteres especiales, números y palabras con menos de tres caracteres, de esta manera lograron aumentar la certeza del análisis.

En el campo de los videojuegos, es raro encontrar un estudio de esta naturaleza, normalmente van enfocados al aspecto psicológico o al desarrollo. Lo que se ha realizado son estudios en base a las reseñas de los juegos, que se hacen tanto de manera profesional como de manera casual (las escritas por los jugadores). En el artículo de (Strååt & Verhagen, 2017) utilizan el análisis de sentimientos para clasificar las reseñas en malas, neutrales o buenas con base a 3 características de los juegos, las cuales son la historia, el combate y los personajes. Dentro de los resultados encontraron que la mayoría de las críticas se enfocaron en la historia, los resultados muestran que la mayoría de las reseñas están enfocadas en la historia y están reseñadas como malas historias.

2.1 Metodología

Dado el amplio campo de aplicación del análisis de sentimientos, se realizan estudios en base a las redes sociales y a las opiniones de los consumidores, (Fang & Zhan, 2015) realizó un estudio utilizando las opiniones de los clientes en la plataforma de Amazon, se utiliza una metodología de 3 fases (Recolección, Identificación de sentimientos y resultados), y al terminar se aplican algunos algoritmos para vislumbrar los principales comportamientos de las reseñas que los clientes han publicado.

Debido a que el SA es un proceso de clasificación, se deben seguir una serie de pasos para realizarlo con éxito, algunos autores manejan metodologías diferentes pero con bases similares, en el caso de (Medhat, Hassan, & Korashy, 2014) mencionan 5 pasos que incluyen la revisión del material, la identificación de los sentimientos, la selección de características, la clasificación de sentimientos y los resultados. En este trabajo, se han realizado adecuaciones a esta metodología para enfocarla a las características del caso de estudio, las cuales se muestran en la figura 1.

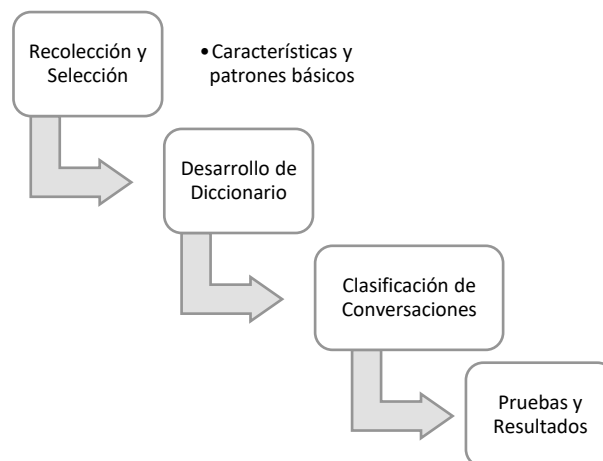


Figura 1. Metodología para el SA basado en el modelo que presenta (Medhat, Hassan, & Korashy, 2014)

La primera etapa consiste en la extracción de la información, seleccionando aquellas muestras útiles para el estudio, seguido del uso de un software para desarrollar el diccionario de palabras en el que se añade un peso informacional a los datos con base al tema o evento. Una vez que el software realiza la evaluación de la información, esta se clasifica en negativo, positivo o neutral y con los resultados se evalúan los aciertos obtenidos de la misma y con ello se procede a experimentar con otros valores para mejorar su precisión.

2.2 Herramientas

Existen diferentes herramientas para el análisis de sentimientos, en muchas ocasiones se presentan como herramientas enfocadas en el aspecto mercantil, es decir, en base a una marca se evalúan los sentimientos u opinión de la misma en determinadas redes sociales y no permiten la modificación de ciertos criterios, como el umbral de decisión o el diccionario que utilizan, un ejemplo de este tipo de sistemas es Brandwatch, una aplicación que permite el análisis de sentimientos de una o varias marcas, comparando las opiniones en redes sociales y así saber que marca es del agrado del público en general, usualmente estas redes sociales son Twitter o Facebook (Brandwatch, s.f.).

Por otro lado existen APIS para generar el análisis de sentimientos, particularmente usando Excel, dos de los conocidos son Meaningcloud y Lexalytics, tienen características similares, dentro de las cosas que lo diferencian es que en Meaningcloud maneja una gran variedad de idiomas en comparación de la API de Lexalytics, también contiene un número mayor de posibles clasificaciones, las cuales son muy positivo y muy negativo, sin embargo, no permite la modificación del umbral para la clasificación y en los resultados no se muestra el valor numérico (Meaningcloud, s.f.).

Lexalytics tiene a disposición 3 APIS para el estudio de documentos: API en la nube, Semantria for Excel (extensión de Excel) y AI Assembler™, que es una herramienta de modelado de máquinas de aprendizaje (Lexalytics, s.f.).

El software que se utiliza para el análisis de sentimientos en este trabajo será Semantria for Excel, tiene la ventaja de reconocer la importancia de una palabra en un enunciado, evaluando el sentimiento a partir de ellas, es decir que, analiza el contexto de las oraciones en relación al contenido y compara los resultados obtenidos. Por otro lado, permite la modificación total del diccionario de palabras, también posibilita la modificación del umbral para la clasificación, con ello se pueden hacer diferentes experimentos buscando una mejor precisión en la clasificación, es bastante accesible ya que trabaja con Excel y no requiere instancias o programas adicionales utilizando un entorno bastante amigable para operar.

3 Análisis de sentimientos en videojuegos

Como se mencionó, el análisis de sentimientos se enfoca en frases que se clasifican como negativas, positivas o neutrales, se pueden añadir dos evaluaciones más que son, Muy Negativo o Muy Positivo, para el caso de este estudio, se considera cada oración de una conversación como una frase, siendo así un proceso similar a los estudios como SemEval-2016 Task 4. Para lograr

determinar la clasificación de cada conversación, se utiliza un promedio que depende de los valores individuales de las oraciones.

A continuación, se presenta lo realizado en cada etapa de la metodología para realizar el análisis de las conversaciones de los videojuegos LOL, HOTS y OVW.

3.1 Recolección y selección

Para la etapa de recolección de la información, además de las conversaciones se incluyen variables como: el nombre del juego, el idioma, la duración, el tipo de partida, el rango, el resultado obtenido y el género del videojuego, las cuales se explican brevemente a continuación.

- Nombre del juego: se refiere al juego del cual se extrajo la conversación que como se indicó puede ser LOL, OVW y HOTS.
- Idioma: se refiere el idioma predilecto de las conversaciones en la partida.
- Duración: se refiere al tiempo total que duro la partida.
- Rango: en los videojuegos se clasifica a los jugadores en base a su destreza, la clase cambia de juego en juego, por ejemplo, mientras en LOL el nivel más alto es Challenger, en HOTS es Grand master y en OVW es Top 500, para uniformar los datos, en este trabajo se ocuparon los rangos bronce, plata, oro, platino y diamante, siendo bronce el más bajo y diamante el más alto.
- Tipo de Partida: se refiere a diferentes modos de juego, en general se divide en dos, competitiva en donde el resultado de la partida afecta el rango del jugador y casual donde se adquiere experiencia para subir de nivel.
- Resultado Obtenido: se refiere a si la partida concluyó con una victoria o derrota, en LOL se permite la rendición, sin embargo, todos los miembros del equipo deben estar de acuerdo y aún así, se toma como derrota.
- Género: se refiere al tipo de juego, por ejemplo, OVW es un juego FPS, por sus siglas en inglés First Person Shooter, traducido al español como juego de disparos en primera persona y en caso de HOTS y LOL son tipo MOBA por sus siglas Multiplayer Online Battle Arena, si bien, no tiene una traducción exacta, es un juego donde se unen equipos de 5 jugadores para destruir una base enemiga.

Dentro de cada juego, se tiene una interfaz para comunicarse, esta interfaz refleja la conversación de los jugadores de un mismo equipo. Para poder recopilar la información existen dos formas, la primera se logra jugando el videojuego, ya que la partida se almacena junto con las conversaciones en un formato de video. Sin embargo, el video únicamente muestra la conversación entre los integrantes del equipo y no la del equipo rival. La segunda forma es a través de plataformas de contenido multimedia donde los jugadores comparten sus partidas, desafortunadamente, en ocasiones llegan a ser editadas por el jugador por lo que no serían útiles para este trabajo.

En ambos casos, extraer las conversaciones es un proceso tardado ya que requiere, en el caso de que se juegue, el doble del tiempo que dure la partida; por ejemplo, una partida tradicional de LOL dura aproximadamente 45 minutos, más otros 45 minutos para repetirla y capturar el texto. En caso de los videos, solo se requiere acceder a la plataforma y reproducirlos para extraer las conversaciones completas.

Para este trabajo en particular se extrajeron 75 conversaciones, de las cuales, 25 están en idioma español, y 50 en inglés; las conversaciones son de longitud variable, la más extensa contiene 440 palabras, mientras que la más pequeña contiene solo 50 palabras. La longitud de una conversación no influye en su clasificación, pueden existir conversaciones muy largas que sean neutrales y otras que en solo 2 oraciones llegan a ser insultantes o negativas. En lo que respecta a la selección, únicamente se consideraron aquellas conversaciones que tuvieran al menos 50 palabras, como resultado tenemos que en LOL se tienen 3935 palabras que se juntan en 1182 oraciones, para HOTS contiene 1431 palabras con 473 oraciones y para OVW contiene 1681 Palabras con 496 Oraciones, donde cada oración representa a un jugador. Las oraciones tienen entre 1 a 15 palabras en promedio.

Como parte fundamental, la información se presenta de manera anónima, dentro de las conversaciones se vislumbran diferentes patrones de manera inmediata:

- Cada línea es una oración escrita por un jugador, es raro el jugador que escribe más de una línea.
- Se llegan a presentar errores ortográficos y tipográficos.
- En el caso de partidas casuales, la comunicación empieza cuando la partida se empieza a complicar, mientras en las competitivas la comunicación se da desde el inicio.
- Aunque cada jugador tiene un alias, es raro cuando se llegan a mencionar, normalmente se ocupan los nombres de los personajes del juego.
- Especialmente en LOL, para escribir palabras obscenas sin ser censurados, los jugadores optan por cambiar algunas letras por asteriscos.
- Si bien no existen “emojis” en estos juegos, se presentan algunas combinaciones de letras para representarlos.
- Aunque algunas conversaciones son en español, sin embargo, se llega a ocupar palabras o acrónimos en inglés.

Para poder realizar el SA se necesita una organización en las conversaciones, cada videojuego cuenta con algunas palabras o acrónimos específicos, algunas de estas palabras tienen o carecen de un significado en los demás juegos, un ejemplo directo son los personajes de cada juego, un personaje de LOL no existe en HOTS, observando esta característica se han organizado todas las conversaciones con

sus respectivos juegos, de esta manera será más sencillo realizar el diccionario de palabras para cada uno y tener una mejor precisión al realizar el análisis.

3.2 Diccionario

Como se explicó en el punto anterior cada juego se tiene que analizar por separado, sin embargo, así como hay palabras específicas para cada juego, también se tienen palabras con un significado similar sin importar el juego o el idioma, el ejemplo más grande que se tiene es el acrónimo “gg”, que se refiere a las palabras “good game”, y se interpreta como un agradecimiento a los jugadores, es algo similar a un apretón de manos al terminar un partido de fútbol.

Aunque el acrónimo “gg” puede llegar a significar algo positivo, también se puede convertir en algo insultante, si se le añaden las letras “ez”, ya que se convierte en algo totalmente diferente. “ez” hace referencia a la palabra en inglés “easy” debido a su similitud al momento de pronunciarse, “ggez” le da a conocer al enemigo que no fue un rival digno, que las habilidades que mostraron no fueron suficientes para combatir y que fue demasiado fácil alzarse con la victoria, en otras palabras alardear de una supuesta superioridad al equipo enemigo.

Para este trabajo, se agregaron palabras relacionadas con los nombres de los personajes de cada juego, para OVW se añadieron 30 palabras, para HOTS 86 y para LOL 145. Además se insertaron palabras y frases de uso común en los 3 juegos; se añadieron 27 acrónimos como “gg” y por último, se recolectaron todas aquellas palabras o frases que son insultantes o agresivas, entre estas, groserías de uso coloquial, combinación de palabras con doble sentido. También se incluyó una lista negra (“blacklist”) con palabras que incluyen temas como homofobia u odio racial.

Adicional al análisis de sentimientos, Semantria for Excel otorga gráficos en donde representa la frecuencia de las palabras en las conversaciones y su peso informacional en el diccionario, como se aprecia en la tabla 1 y figura 2, existen palabras con un mayor tamaño, esto quiere decir que esa palabra fue mayormente ocupada, el color de la palabra es distintivo de su peso informacional, desde rojo que representa lo negativo, hasta verde que indica palabras positivas, pasando por una escala de grises que son las palabras neutrales que contienen un valor diferente a cero.

Phrase	negative	neutral	positive	Total general
xd	-	23	-	23
gg	-	15	-	15
bien	-	6	1	7
puto	6	-	-	6
vi	-	6	-	6
pendejo	5	-	-	5
bueno	-	5	-	5
:d	-	5	-	5
lux	-	5	-	5
ashe	-	4	-	4
supp	-	4	-	4
mala	4	-	-	4
daño	-	4	-	4
jinx	-	4	-	4
perdón	-	4	-	4
zed	-	4	-	4
buena	-	4	-	4
Total general	15	93	1	109

Tabla 1. Frecuencia de las palabras en las conversaciones de LOL



Figura 2. Representación gráfica de la frecuencia en tabla 1

3.3 Clasificación

Para el análisis de sentimientos, Semantria for Excel busca las oraciones en el diccionario de palabras y añade su valor individual para calcular el valor de toda la conversación y asignarle una clase. Un ejemplo de esto se muestra en la figura 3, en la cual se presenta una oración con las palabras “jax” y “mid” las cuales no parecen tener mucho sentido en la frase.

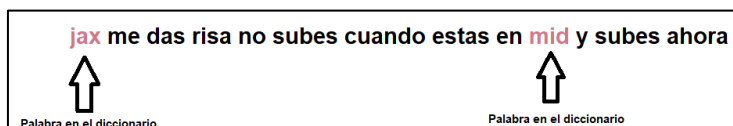


Figura 3. Ejemplo de una oración de LOL

La primera palabra refiere a un personaje del juego “jax” y la segunda a un espacio del mapa como se muestra en la figura 4. En el juego LOL existen 3 líneas donde los personajes pueden transitar para lograr llegar a la base enemiga, la parte de arriba se le conoce como “top”, la de en medio como “mid” (que es la segunda palabra) y la de abajo “bot”; con este conocimiento la oración entonces ya tiene sentido, pues se refiere a que el jugador que maneja al personaje Jax, no trató de llegar a la base enemiga cuando estaba en una posición más sencilla que la actual.



Figura 4. Mapa de LOL. Fuente: (Riot, 2009)

Una vez se tengan las conversaciones en un documento de Excel, se puede generar el análisis. En la barra de herramientas de Lexalytics se encuentra un botón el cual lleva por nombre Start (figura 5), al seleccionar esta opción aparecerá una ventana emergente la cual solicita seleccionar el documento con el texto a analizar (figura 6), Semantria solicitará la identificación de las columnas, cuál es el ID, la conversación y los datos extras que pueda contener el documento (figura 7). Posteriormente, se selecciona cuántas filas se van a considerar, pueden ser todas o un intervalo en específico (figura 8), y para concluir, se le da un nombre al análisis, alguna descripción y el diccionario que se ocupará (figura 9).

El proceso se inicia y cuando culmine nos mostrará en color verde los procesos realizados (figura 10).

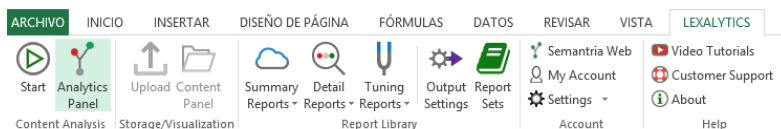


Figura 5. Barra de Herramientas de Lexalytics, fuente: (Lexalytics, s.f.)

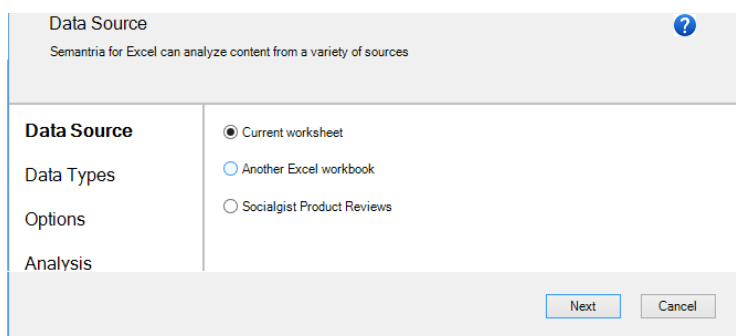


Figura 6. Elección de archivo a evaluar, fuente: (Lexalytics, s.f.)

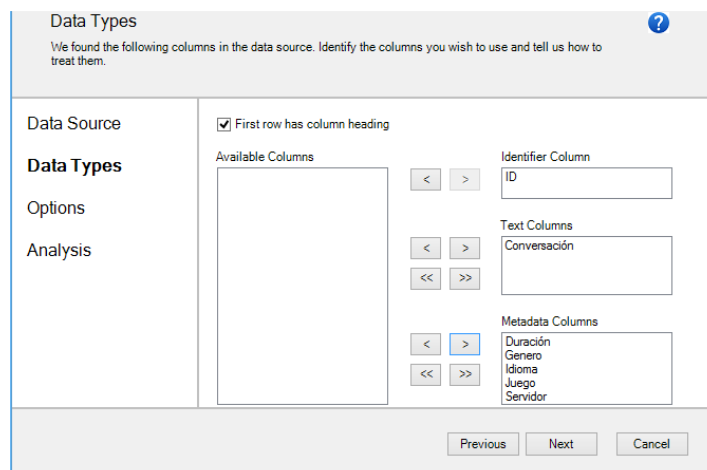


Figura 7. Selección de Información fuente: (Lexalytics, s.f.)

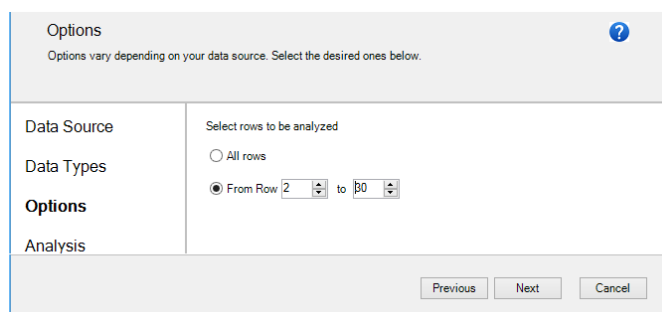


Figura 8. Tamaño de filas a considerar, fuente: (Lexalytics, s.f.)

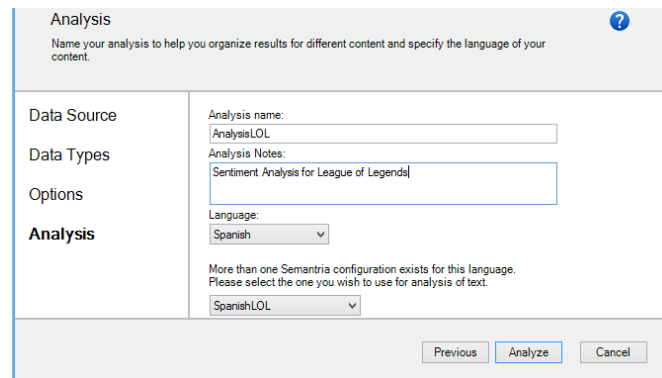


Figura 9. Nombre, Descripción y Selección del Diccionario para el Análisis, fuente: (Lexalytics, s.f.)

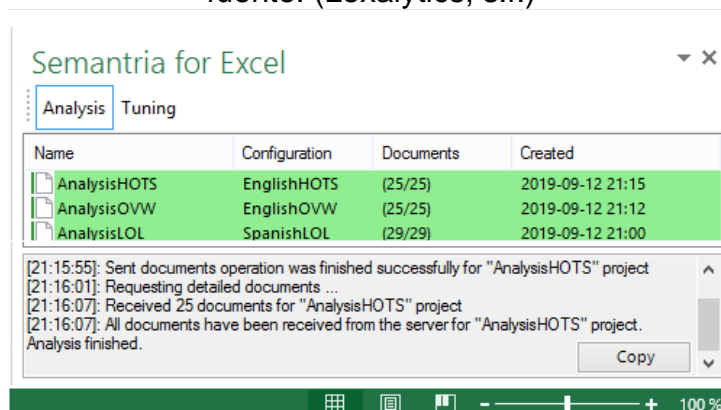


Figura 10. Análisis realizados en Semantria, fuente: (Lexalytics, s.f.)

Con base a la documentación proporcionada por Lexalytics para clasificar las conversaciones, Semantria for Excel trabaja de la siguiente forma, se utilizan 3 variables base:

P = palabra

F = frase

N = palabra o frase neutral

Los valores que P o F pueden obtener están en el rango: $(-\infty, -0.05]$ v $[0.2, +\infty)$, mientras que para N, el rango es $(-0.05, 0.2)$.

De esta manera, las palabras o frases serán clasificadas como NEGATIVAS, NEUTRALES y POSITIVAS: Si P o F cae en el rango $(-\infty, -0.05]$ se evalúa como una frase o palabra NEGATIVA, mientras que si cae en el rango $[0.2, +\infty)$ se evalúa como POSITIVA y finalmente si cae en el rango $(-0.05, 0.2)$ se evalúa como NEUTRAL.

Para calcular el valor promedio de cada conversación se tiene la ecuación 1:

$$CN = \left(\frac{\sum_{i=1}^{i=n} P_i}{n} + \frac{\sum_{i=1}^{i=n} F_i}{n} + \frac{\sum_{i=1}^{i=n} N_i}{n} \right) / 3 \quad (1)$$

Donde

P_i es el valor de la palabra en el diccionario.

F_i es el valor de la frase en el diccionario

N_i es el valor de las palabras o frases neutrales en el diccionario

Además se deben cumplir las siguientes condiciones:

$$\begin{aligned} P_i &\neq P_{i+1} \\ F_i &\neq F_{i+1} \\ N_i &\neq 0 \end{aligned}$$

En términos generales lo que realiza la fórmula y por consiguiente Semantria for Excel, es realizar la sumatoria de todas las palabras o frases que encuentre en la conversación y al final dividirla entre el número total de palabras o frases. Una vez calculado CN, se calcula TCN, como el valor promedio de todas las conversaciones (ver ecuación 2)

$$TCN = \frac{\sum_{i=1}^{i=NP} CN_i}{NP} \quad (2)$$

Donde:

CN es el valor de la conversación de cada partida y NP el número total de partidas.

El umbral que se utiliza para la clasificación es el recomendado por Semantria for Excel el cual proporciona un punto de partida para determinar si una conversación es positiva o negativa. Dicho umbral se puede modificar con base a las preferencias del usuario y a los patrones que se puedan observar.

3.4 Pruebas y resultados

El umbral utilizado para la clasificación de las conversaciones es el dado por defecto en el API de Semantría for Excel. Los resultados se pueden observar en la tabla 2.

Juego	OVW	LOL	HOTS
Conversaciones Positivas	3	1	7
Conversaciones Negativas	11	9	7
Conversaciones Neutrales	11	15	11
Valor promedio de las Conversaciones	-0.0077	-0.0205	0.0032
Promedio de duración	14.28 min	36.9 min	20.45 min
Victorias totales	10	14	7

Tabla 2. Resultados del análisis con Semantria for Excel en todas las partidas

En la tabla 2, se aprecia que la cantidad de conversaciones negativas prevalece en los 3 juegos, sin embargo en caso particular de LOL, solo existe 1 positiva; por otro lado, aunque HOTS tiene 11 negativas, es el que tiene la mayor cantidad de conversaciones positivas. En el caso de OVW se observa que existe un equilibrio entre las conversaciones positivas y negativas. Al momento de observar el valor promedio de las conversaciones, que nos refleja que tan Negativas o Positivas fueron las conversaciones se puede ver que aunque LOL no tuvo la mayor cantidad de conversaciones negativas, fue el que tuvo la evaluación más baja, lo que significa que tuvo las conversaciones más negativas. De manera general, en los 3 juegos las conversaciones neutrales son mayoría.

Con el fin de evaluar la certeza del software, se agregó la evaluación de un experto que con base al contenido de las conversaciones las clasificó según su criterio, posteriormente se compararon con los datos ofrecidos por el software de la clasificación dada a la conversación por parte del jugador. De esta forma, la clasificación del sistema obtuvo un 62.6% de certeza, al clasificar de manera correcta 47 de las 75 conversaciones. Para determinar si se podía obtener una mejor precisión en la clasificación, se experimentó utilizando diferentes umbrales para clasificar las conversaciones. En la tabla 3 se muestran los resultados obtenidos.

Umbral	Negativa	Neutral	Positiva	Acierto
-.2 a .2	8	55	12	47
-.5 a .5	3	70	2	41
-.05 a .2	27	37	11	47
-.1 a .1	19	33	23	42
-.05 a .05	27	20	28	36
-.1 a 0.5	19	28	28	37
Experto	18	45	12	75

Tabla 3. Resultados de los experimentos

En la tabla 3, se observa que modificar el umbral no aumenta la precisión, el único experimento que dio resultados similares fue el del primer renglón, en el modificar el umbral se incrementó el número de conversaciones neutrales.

4 Aplicación del algoritmo J-48

El algoritmo J-48 tiene una estructura de un árbol de decisión, para lograr realizar la clasificación obtiene un serie de condiciones o reglas con una jerarquía previamente dada, esta jerarquía dependerá del árbol de decisión que se esté ocupando, en base a las variables de entrada devolverá una respuesta para el problema planteado, cada nodo tendrá una salida, estas se dividen en dos partes, las cuales son progresivas o son aquellas que finalizan la clasificación (Aggarwal C., 2015).

En el caso específico del algoritmo J-48, es un algoritmo de inducción, donde se crean subconjuntos de los datos de entrenamiento, en otras palabras, particiones de los datos, en base a estas particiones se generan las reglas pertinentes para lograr clasificar correctamente a el mayor número de casos (Suca, Córdova, Condori, Cayra, & Sulla, 2016).

La precisión del algoritmo va a variar dependiendo del número de particiones que se tengan, aunque en general utiliza el 80% de estos valores como entrenamiento para después probar con el 20% restante, una vez realizada la primer iteración, intercalará entre las particiones, esto quiere decir que hará una combinación de todas las particiones para lograr entrenar y probar entre todas para generar una serie de reglas o condiciones más precisas (Aggarwal C., 2015).

Para aplicar el algoritmo en este caso se añadieron las variables: Nombre del juego, Idioma, Duración de la Partida, Rango del jugador, Tipo de Partida, Resultado Obtenido, Género y la evaluación de las conversaciones, el algoritmo J-48 mostrará, que variable es la más relevante dentro de la clasificación de las

conversaciones, en otras palabras, averiguar si el ambiente (las variables) afecta de manera significativa a la naturaleza de las conversaciones.

En la figura 11, se aprecia los atributos que fueron tomados en cuenta para la aplicación del estudio, las variables que fueron relevantes fueron la puntuación que es la evaluación dada por Semantria for Excel siendo esta variable la indispensable para la clasificación, posteriormente se aprecia que el resultado de la partida infiere en caso de que la evaluación sea muy negativa.

Como se aprecia en la figura 12 la precisión que muestra el algoritmo es de 61.33%, en comparación de la dada por Semantria for Excel no se observa un aumento, esto debido que el algoritmo toma como variable principal las conversaciones, adicional, se muestra la matriz de confusión del algoritmo, el comportamiento que ofrece es el clasificar en gran medida las conversaciones en Neutral debido a la gran cantidad que existen, en la figura 13 se observa el árbol del algoritmo mostrando de manera gráfica este comportamiento.

```
Classifier output
=== Run information ===
Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    datos generales-weka.filters.unsupervised.attribute.Remove-R1
Instances:   75
Attributes:  8
             Juego
             Punt
             Duración
             Resultado
             Tipo de Partida
             Genero
             Rank
             Clasificación
Test mode:   5-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree
-----
Punt <= -0.076167
| Resultado = V: Neutral (8.0/3.0)
| Resultado = D: Negative (17.0/5.0)
Punt > -0.076167
| Punt <= 0.2435: Neutral (41.0/10.0)
| Punt > 0.2435: Positive (9.0/4.0)
```

Figura 11. Atributos y comportamiento del algoritmo J-48

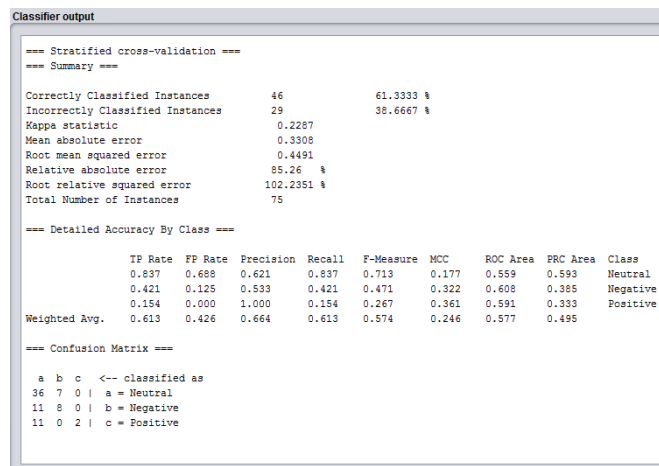


Figura 12. Precisión en la clasificación y matriz de confusión del algoritmo J-48

Al excluir la variable de puntuación (Punt), que es la dada por Semantria, observamos un comportamiento en donde las variables principales que se toman en consideración son el resultado del juego (Victoria o Derrota), en caso de derrota verifica el juego (OVW, HOTS o LOL) y concluye con revisar el tipo de partida o la duración, este comportamiento se puede apreciar en la figura 14.

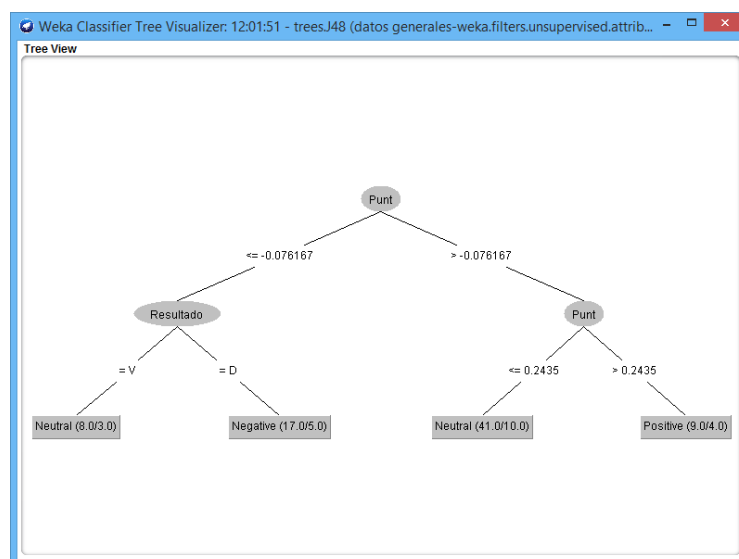


Figura 13. Árbol de clasificación J-48

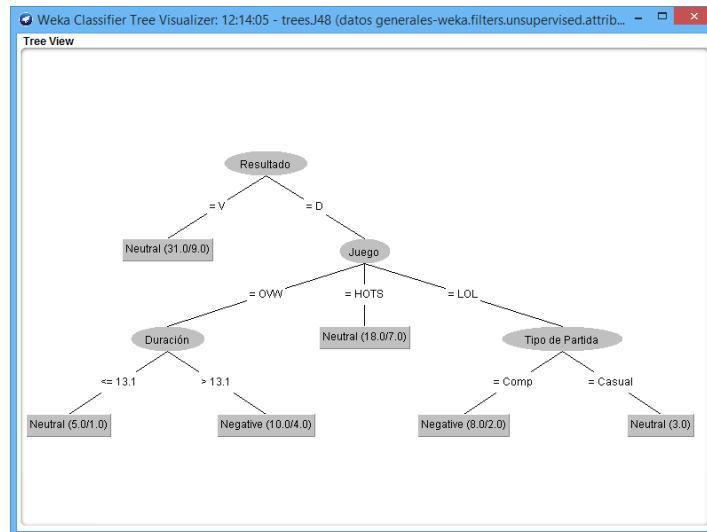


Figura 14. Comportamiento del algoritmo sin la variable puntuación

5 Conclusiones y trabajo futuro

En este artículo se presentaron los resultados de aplicar el análisis de sentimientos a 75 conversaciones de longitud variable escritas en inglés y español, extraídas de partidas de los videojuegos *League of Legends*, *Overwatch* y *Heroes of the Storm*. Utilizando Semantría para Excel, un API del software Lexalytics, se creó un diccionario que pudiera ser utilizado para la clasificación de las conversaciones. Las clases se definieron como positivas, negativas o neutrales.

Por medio de los experimentos realizados se observa que sí es posible aplicar el análisis de sentimientos en las conversaciones de una partida de videojuegos, obteniendo un 62.6% de certeza en la clasificación. Aunque el resultado puede considerarse bajo, se justifica por el desbalance en el número de conversaciones de cada clase, es necesario tener una cantidad equitativa de los tres tipos de conversaciones (Negativa, Positiva y Neutral) esperando obtener un mayor grado de certeza.

Si bien existe un desbalance en las conversaciones, se atribuye a que en gran medida las conversaciones tienden a ser negativas reflejando un comportamiento dentro de las 3 comunidades (*Overwatch*, *League of Legends* y *Heroes of the Storm*), en general se aprecia que así como en otros deportes, las riñas llegan a presentarse en cualquier nivel de juego, siguiendo el mismo concepto que se ha manejado siempre, la competitividad.

Por otro lado el algoritmo C4.5 o J-48 mostró la naturaleza del comportamiento de las clasificaciones y ratificó que en esencia los implicados en las conversaciones tienen mucho que ver en sus comportamientos Negativos o Positivos.

Como trabajo futuro, se considera contar con un mayor número de conversaciones, resolver el problema del desbalance, así como extender de tres a cinco, las clases para mejorar la clasificación incluyendo la actualización del diccionario.

Se considera hacer un análisis comparativo de algoritmos de clasificación con la finalidad de observar si se tienen mejores resultados que el algoritmo C4.5 que permita aumentar la certeza de la clasificación.

REFERENCIAS

- Aggarwal C., C. (2015). *Data Mining*. Switzerland: Springer. doi:10.1007/978-1-4614-3223-4
- Brandwatch. (s.f.). *Brandwatch Analytics*. Recuperado el 16 de 04 de 2019, de <https://www.brandwatch.com/es/products/analytics/>
- Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, 2, 14. doi: 10.1186/s40537-015-0015-2
- Harris, B. (2014). *Console Wars: Sega Vs Nintendo-and the Battle that Defined a Generation*. Atlantic Books Ltd.
- Lexalytics. (s.f.). *Data Analytics with NLP & Text Analytics*. Recuperado el 20 de 04 de 2019, de <https://www.lexalytics.com/technology/summarization>
- Mat Zin, H., Mustapha, N., Azmi Murad, M., & Mohd Sharef, N. (2017). The Effects of Pre-Processing Strategies in Sentiment. *AIP conference proceedings*, 20-89. doi:10.1063/1.5005422
- Meaningcloud. (s.f.). *Meaningcloud Sentiment Analysis in Excel*. Recuperado el 03 de 04 de 2019, de <https://www.meaningcloud.com/developer/excel-addin/doc/3.4/sentiment-analysis>
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications. A survey. *Ain Shams engineering journal*, 5, 1093-1113. doi:10.1016/j.asej.2014.04.011
- Navok, P., Ritter, A., Rosenthal, S., Sebastiani, F., & Stoyanov, V. (2016). SemEval-2016 Task 4: Sentiment Analysis in Twitter. *Proceedings of the 10th international workshop on semantic evaluation (semeval-2016)*, 1-18. doi:10.18653/v1/S16-1
- Riot, G. (27 de Octubre de 2009). League of Legends. USA. Obtenido de <https://lan.leagueoflegends.com/es/game-info/>

Rosenthal, S., Farra, N., & Nakov, P. (2017). SemEval-2017 Task 4: Sentiment Analysis in Twitter. *Proceedings of the 11th international workshop on semantic evaluation*, 502-518. doi:10.18653/v1/S17-2088

Steven, K. (2016). *The Ultimate History of VideoGames*. B DE BOOKS.

Strååt, B., & Verhagen, H. (2017). Using User Created Game Reviews for Sentiment. *GHITALY@CHIItaly*.

Suca, C., Córdova, A., Condori, A., Cayra, J., & Sulla, J. (2016). Comparación de algoritmos de clasificación para predicción de casos de obesidad infantil. *Perú: Universidad Nacional de San Agustín*.

Waikato, U. d. (31 de 10 de 2011). Weka. Nueva Zelanda: GNU-GPL.

Notas biográficas



Daniel Iván Ángeles Gómez Recibió el certificado total de estudios de la carrera de Ingeniería en Sistemas y Comunicaciones en la Universidad Autónoma del Estado de México. Actualmente, se encuentra en la búsqueda de su titulación y crecimiento profesional. Sus áreas de interés son el desarrollo de software, análisis de datos y desarrollo de videojuegos.



Dra. Maricela Quintana López es Profesora de Tiempo Completo del Centro Universitario UAEM Valle de México de la Universidad Autónoma del Estado de México. La Dra. Quintana es Ingeniera en Sistemas Computacionales, y tiene la Maestría y el Doctorado en Ciencias Computacionales con especialidad en Inteligencia Artificial por el Tecnológico de Monterrey. Cuenta con el reconocimiento de Perfil Deseable PRODEP de la SEP en México. Ha dirigido tesis de maestría y licenciatura, ha sido conferencista magistral y ponente en diversos eventos. Ha publicado artículos en las áreas de Inteligencia Artificial y Minería de Datos relacionados con sus proyectos de Investigación.



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 2.5 México.