



**UNIVERSIDAD DE GUANAJUATO**

---

---

**CAMPUS IRAPUATO – SALAMANCA**

**DIVISIÓN DE INGENIERÍAS**

**SISTEMA AUTOMÁTICO DE CLASIFICACIÓN DE LA  
POLARIDAD DE OPINIONES DE PELÍCULAS DE  
LARGOMETRAJE**

**TESIS**

**QUE PARA OBTENER EL GRADO DE:  
MAESTRÍA EN ADMINISTRACIÓN DE TECNOLOGÍAS**

**PRESENTA:**

**Ing. Jesús Andrés Sierra Rangel**

**DIRECTORES DE TESIS:**

**Dr. Rafael Guzmán Cabrera**

**Dr. Everardo Vargas Rodríguez**



---

# CONTENIDO

---

|   |     |
|---|-----|
| Lista de tablas .....                               | IV  |
| Lista de figuras .....                              | V   |
| Agradecimientos.....                                | VII |
| 1 Resumen.....                                      | 1   |
| 2 Introducción.....                                 | 2   |
| 2.1 Aplicaciones del análisis de sentimientos ..... | 4   |
| 2.1.1 Salud mental.....                             | 4   |
| 2.1.2 Ciencias sociales.....                        | 4   |
| 2.1.3 Películas.....                                | 4   |
| 2.2 Estado del Arte.....                            | 6   |
| 2.3 Large movie review dataset .....                | 8   |
| 2.4 Justificación del proyecto de tesis.....        | 9   |
| 2.5 Preguntas de investigación .....                | 10  |
| 2.6 Hipótesis .....                                 | 10  |
| 2.7 Objetivos .....                                 | 11  |
| 2.7.1 Objetivo General.....                         | 11  |
| 2.7.2 Objetivo Específico.....                      | 11  |
| 3 Marco Teórico .....                               | 12  |
| 3.1 ¿Qué es la Inteligencia Artificial? .....       | 12  |
| 3.2 Aprendizaje Automático .....                    | 13  |
| 3.3 Tipos de aprendizaje de máquina .....           | 14  |
| 3.3.1 Aprendizaje Supervisado.....                  | 14  |
| 3.3.2 Aprendizaje no supervisado .....              | 15  |
| 3.4 Minería de datos .....                          | 15  |

|       |   |    |
|-------|---|----|
| 3.4.1 | Minería de opiniones .....                | 16 |
| 3.5   | Diccionarios léxicos.....                 | 18 |
| 3.6   | Análisis de Sentimientos .....            | 19 |
| 3.7   | Orientación semántica .....               | 22 |
| 3.8   | Modelos de aprendizaje automático.....    | 22 |
| 3.8.1 | Naïve Bayes .....                         | 22 |
| 3.8.2 | Regresión Logística.....                  | 28 |
| 3.8.3 | Máquinas de soporte vectorial.....        | 31 |
| 3.8.4 | Método de K-Vecinos más cercanos .....    | 32 |
| 3.8.5 | Bosque Aleatorio .....                    | 34 |
| 4     | Metodología .....                         | 36 |
| 4.1   | Entrada de datos.....                     | 37 |
| 4.2   | Selección de datos.....                   | 38 |
| 4.3   | Tokenizar .....                           | 38 |
| 4.4   | Preprocesamiento.....                     | 39 |
| 4.4.1 | Normalización.....                        | 40 |
| 4.4.2 | Eliminación de acentos.....               | 40 |
| 4.4.3 | Lematización .....                        | 41 |
| 4.4.4 | Eliminación de mayúsculas .....           | 41 |
| 4.4.5 | Eliminación de palabras de paro .....     | 41 |
| 4.5   | Bolsa de palabras .....                   | 42 |
| 4.6   | Escenario de clasificación .....          | 43 |
| 4.7   | Modelos del aprendizaje automático.....   | 43 |
| 4.8   | Evaluación del Desempeño del Sistema..... | 43 |
| 4.9   | Predicción .....                          | 45 |
| 5     | Implementación del Clasificador .....     | 46 |
| 5.1   | Conjunto de datos .....                   | 46 |
| 5.2   | Configuración .....                       | 47 |
| 5.2.1 | Punto de partida (baseline).....          | 48 |

|       |  |    |
|-------|--|----|
| 5.3   | Proceso.....   | 48 |
| 5.3.1 | Proceso de la entrada de datos a creación de bolsa de palabras ..... | 48 |
| 5.3.2 | Análisis de sentimiento.....   | 50 |
| 5.3.3 | Predicción.....  | 53 |
| 5.4   | Resultados del Baseline.....   | 54 |
| 5.4.1 | Punto de partida .....   | 54 |
| 6     | Resultados.....  | 56 |
| 6.1   | Exactitud considerando diferentes tipos de pre-procesamiento.....    | 56 |
| 6.2   | Precisión considerando diferentes tipos de pre-procesamiento.....    | 58 |
| 6.3   | Recall considerando diferentes tipos de pre-procesamiento.....       | 59 |
| 6.4   | F1-score considerando diferentes tipos de pre-procesamiento .....    | 61 |
| 6.5   | Resultado final .....  | 62 |
| 7     | Conclusiones.....  | 64 |
| 8     | Producción científica .....  | 66 |
|       | Referencias .....  | 69 |
|       | Apéndice A.....  | 80 |
|       | Apéndice B.....  | 84 |

---

## Lista de tablas

---

|   |    |
|---|----|
| Tabla 1. Ejemplo de datos con etiqueta para Naive Bayes .....                     | 24 |
| Tabla 2. Punto de partida con métricas .....                                      | 55 |
| Tabla 3. Exactitud de los 4 tipos de pre-procesamiento .....                      | 57 |
| Tabla 4. Precisión alcanzada considerando diferentes tipos pre-procesamiento ..   | 58 |
| Tabla 5. Recall obtenido al aplicar tipos de pre-procesamiento .....              | 60 |
| Tabla 6. F1-score obtenidos al aplicar diferentes tipos de pre-procesamiento..... | 61 |
| Tabla 7. Mejor resultado con métricas.....  | 63 |

---

## Lista de figuras

---

|   |    |
|---|----|
| Figura 1. Ejemplo de una quintupla en un análisis de Alcalde, 2021. ....                  | 20 |
| Figura 2. Ejemplo de conjunto de datos para Naive Bayes .....                             | 23 |
| Figura 3. Ejemplo con histograma y probabilidades .....                                   | 25 |
| Figura 4. Ejemplo de probabilidad para cada conjunto .....                                | 26 |
| Figura 5. Probabilidad de X en el ejemplo.....  | 26 |
| Figura 6. Sustitución de los valores en ecuación .....                                    | 27 |
| Figura 7. Ejemplo para determinar categoría de un nuevo mensaje .....                     | 27 |
| Figura 8. Histograma con adición de un elemento para mejor clasificación .....            | 28 |
| Figura 9. Ejemplo con histograma cambiado para determinar categoría de nuevo mensaje..... | 28 |
| Figura 10. Ejemplo de gráfico de regresión logística.....                                 | 31 |
| Figura 11. Ejemplo de un gráfico por categorías .....                                     | 32 |
| Figura 12. Ejemplo de K-vecinos más cercanos .....  | 33 |
| Figura 13. Ejemplo de un árbol de decisión .....  | 35 |
| Figura 14. Diseño de algoritmo de procesamiento.....                                      | 36 |
| Figura 15. Ejemplo de entrada de datos con salida esperada.....                           | 38 |
| Figura 16. Ejemplo de tokenización .....  | 39 |
| Figura 17. Ejemplo de vector de carácter.....   | 39 |
| Figura 18. Ejemplo sin normalizar .....   | 40 |
| Figura 19. Reducción de matriz.....   | 42 |
| Figura 20. Ejemplo de normalización .....   | 42 |
| Figura 21. Ejemplo de matriz de confusión .....   | 44 |
| Figura 22. Ejemplo de opinión negativa .....  | 46 |
| Figura 23. Ejemplo de opinión positiva.....   | 47 |
| Figura 24. Ejemplo de documento.....  | 48 |
| Figura 25. Función de tokenización.....   | 49 |
| Figura 26. Código para crear matriz documento término. ....                               | 49 |
| Figura 27. Ejemplo de bolsa de trabajo .....  | 50 |
| Figura 28. Ejemplo de bolsa de manera gráfica .....                                       | 50 |
| Figura 29. Conjuntos de datos de análisis.....  | 51 |
| Figura 30. Ejemplo de bolsa de palabras para análisis de sentimiento.....                 | 51 |
| Figura 31. Ejemplo gráfico de bolsa de palabras para el análisis de sentimientos          | 51 |
| Figura 32. Código para separar datos en prueba y entrenamiento. ....                      | 52 |
| Figura 33. Inicialización y entrenamiento de clasificación. ....                          | 52 |
| Figura 34. Ejemplo de Exactitud.....  | 52 |

|  |    |
|--|----|
| Figura 35. Exactitud del ejemplo. ....                                 | 53 |
| Figura 36. Ejemplo de opinión sin etiqueta.....                        | 53 |
| Figura 37. Código para predecir.....                                   | 54 |
| Figura 38. Resultado del sistema. ....                                 | 54 |
| Figura 39. Gráfica de Exactitud para todos los tipos de procesos.....  | 57 |
| Figura 40. Gráfica de precisión para todos los tipos de procesos. .... | 59 |
| Figura 41. Gráfica de recall para todos los tipos de procesos.....     | 60 |
| Figura 42. Gráfica de f1-score para todos los tipos de procesos. ....  | 62 |



---

## Agradecimientos

---

Agradezco a mi familia nuclear por apoyarme en esta decisión de seguir estudiando, a mis compañeros y seres queridos durante este trayecto, a los directores de tesis y mis síndos por el apoyo otorgado, y a esas personas que formaron parte de esta aventura, se los agradezco.

Agradezco al Consejo Nacional de Humanidades Ciencia y Tecnologías (CONAHCYT) por el apoyo brindado a CVU: **1165067** para realizar mi investigación y brindar apoyo a la comunidad investigadora en el programa Becas Nacional (Tradicional) 2022 – 1, especialmente con el apoyo de beca número 917912.

---

## 1 Resumen

---

### **SISTEMA AUTOMÁTICO DE CLASIFICACIÓN DE LA POLARIDAD DE OPINIONES DE PELÍCULAS DE LARGOMETRAJE**

El crecimiento del big data en el internet es una de las áreas de oportunidad que las empresas buscan aprovechar para poder obtener una retroalimentación de los usuarios quienes prefieren dejar una opinión o reseña a realizar una encuesta, es por ello por lo que los últimos años el reto para las compañías es analizar esa información para obtener datos específicos que les brinden una ventaja frente a sus competidores, para lo cual el aprendizaje automático representa una herramienta sumamente útil. En este trabajo de tesis se desarrolló un sistema que contiene 5 modelos de aprendizaje: naive bayes (NB), máquinas de soporte vectorial (MSV), k-vecinos más cercanos(K-VC), regresión logística (RL) y bosque aleatorio (BA), los cuales son sometidos a procesos sin y con etapas de pre-procesamiento para poder determinar las mejores opciones para cada uno. El resultado sorprende por el alto grado de exactitud que se tiene, siendo NB el mejor, logrando un 97.5% de exactitud. Con solo 400 opiniones y se descubrió que es mejor no tener pre-procesamiento para obtener un mejor rendimiento, es de destacar en la búsqueda del estado del arte no se encontró un modelo que pudiera alcanzar una exactitud del 97.5%. Con los demás modelos se logró tener incrementos del 12% en exactitud y precisión con eliminación de acentos y eliminación de palabras de paro para máquinas de soporte vectorial, incremento del 14% para la exactitud y precisión usando eliminación de palabras de paro para bosque aleatorio, en K-Vecinos más cercano se logró un incremento del 8% con eliminación de palabras de paro y regresión logística logró alcanzar una exactitud del 0.925 con eliminación de acentos y eliminación de palabras de paro.

---

## 2 Introducción

---

Actualmente la industria del entretenimiento es uno de los negocios con mayor volumen de ingresos generados a nivel mundial [1], provocando tener una situación social favorable que beneficia este medio desde la llegada de la pandemia, la cual marcó un antes y después para la sociedad mundial [2], generando un cambio en la interacción de los públicos al ser aislados, momento donde se aumentó el número de clientes de las plataformas de “streaming”, lo cual se define como la distribución digital de archivos a través de la difusión continua por una red informática [3]. Esto ha llevado a la creación de sociedades en la red, donde cada persona publica su opinión sobre si le gusta o no el producto, esto sumado a diferentes canales de información, como: redes sociales, canales dedicados, blogs, sitios web de reseñas de películas y series (por ejemplo: Rotten Tomatoes, Internet Movie Database, Filmaffinity, etc.) provocó la creación de enormes cantidades de opiniones sobre un producto o servicio, la explotación de estos datos requiere de herramientas del área de inteligencia artificial, tales que puedan extraerles información [1], [4], [5].

La importancia de esta tendencia puede observarse en el informe de “Global Digital Trends” que indica que el número de usuarios en el planeta para abril del 2023 fue de aproximadamente 5,180 millones de usuarios, lo que representa un 64.6% de la población mundial. Además, los usuarios pasan en promedio de 6 horas con 35 minutos en internet, y de estos el porcentaje de usuarios con edades entre 16 a los 64 años que desean ver algo por medio de “streaming” en TV es del 90.1% [6].

A lo largo de los años se ha establecido el alto grado de causalidad que desempeñan las reseñas y los comentarios en línea sobre el resultado de una película. Frecuentemente, las opiniones tienen una gran influencia debido a que las personas deciden que producto consumir tomando en consideración las evaluaciones y opiniones de una tercera persona [7]. Para el caso de productos como películas, se genera una gran cantidad de datos de forma de “electronic word of mouth” (eWOM) que se refiere a la forma de definir un intercambio de información no comercial escrita. Una de las tareas que en años recientes se ha vuelto popular es el análisis de sentimientos [8], el cual involucra varias áreas como son: el procesamiento del lenguaje natural, el uso de minería de texto y el aprendizaje automático. La actividad de analizar el texto de opiniones o reseñas se realiza con el objetivo de identificar la emoción detrás de la opinión (a favor, en contra, positivo, negativo, etc.). En este sentido, algunos autores definen como minería de opiniones, al proceso de identificar la emoción a partir de miles de comentarios de manera

automática [9]. La revisión de sentimientos es una de las áreas de mayor crecimiento por su interés y un potencial de crecimiento considerable. Con la apertura del internet y la tecnología actual, se ha incrementado la disponibilidad de datos de gran volumen, debido a que cada persona puede publicar su opinión de manera rápida en la red [10]. Cada vez es más común que, los usuarios emitan opiniones tanto en blogs, como en redes sociales lo que hace necesario el desarrollo de nuevas metodologías que permitan acceder, organizar e interpretar esta información. Esto se transforma en un reto para la gestión de los datos masivos en la red, lo cual nos lleva a buscar maneras de comprender la sociedad a través del análisis de los datos [11].

Los datos prometen ser un cambio importante para las estrategias empresariales, siendo esto un nuevo factor a considerar en la implementación de sistemas de tomas de decisiones basadas en datos. Pero si nos adentramos un paso más allá, viéndolo desde un lado más humano, nos permite analizar los datos con base a los sentimientos que aportan y así inferir la emoción de los usuarios sienten al opinar de algo real, lo cual nos lleva a buscar maneras de comprender la sociedad a través del análisis de los datos [11].

El análisis de opiniones es un tema que cada día va tomando más peso debido al creciente uso de las redes sociales, las sociedades en la red y los públicos, que producen y reproducen sus capacidades de creación de contenido, participan de manera activa en las discusiones de temas: sociales, políticos, deportivos, entretenimiento y servicios [4].

En la actualidad muchos trabajos científicos están enfocados en la extracción de las emociones de las personas a partir del análisis de sus opiniones de la red, centrándose en un análisis de sentimiento como metodología basada en la minería de textos sobre el “big data” [2], [12]. El disponer de una fuente tan rica de información representa oportunidades para los interesados, pero también es un tema complejo, porque requiere de un análisis y una confiabilidad que brinden las herramientas dedicadas a solucionar de manera precisa la extracción de las opiniones [1].

El análisis de sentimientos va encaminado a un método automático de clasificación de grandes conjuntos de datos en forma de texto en dos categorías, positivas y negativas. La extracción de información de utilidad para las empresas se encuentra de manera no estructurada, es decir, en forma de opinión [13]. Esto es utilizado por algunas organizaciones para el análisis de su imagen. Existen también otros tipos

de aplicaciones como son: opiniones sobre deportes, turismo, política, educación, salud, finanzas, automóviles y la valoración de películas [14].

## 2.1 Aplicaciones del análisis de sentimientos

### 2.1.1 *Salud mental*

En las áreas de salud, un ejemplo es el de Aguilera, et al [15] en el cual hacen uso de una base de datos de eRisk, y usa el análisis de sentimientos con la finalidad de detectar los desórdenes mentales como anorexia y depresión en redes sociales, los autores realizan la normalización de datos para el preprocesamiento y utilizan los clasificadores de one-class classification (OCC) y K-Vecinos más cercanos para el procesamiento. En otros campos de investigación donde se ha realizado el análisis de sentimientos, se pueden mencionar, como ejemplo: el de Arango Pastrana, et al [16] donde se presenta el análisis de sentimientos aplicado a comentarios sobre el tema del aislamiento que existió en la época de la pandemia, con la finalidad de identificar los temas y palabras más frecuentes utilizados en este contexto.

### 2.1.2 *Ciencias sociales*

En el área de ciencias sociales se pueden encontrar diversos trabajos, por ejemplo, en el trabajo de Oztürk, et al [17] se presenta una propuesta para la clasificación de opiniones que tienen usuarios de Twitter sobre los refugiados sirios en dos idiomas, con la finalidad de tener una cantidad considerable de comentarios. En dicho trabajo, normalizaron los datos para el preprocesamiento y desarrollaron un léxico que otorgaba puntuación a los comentarios de acuerdo con el tipo de sentimiento al que pertenecía cada comentario.

### 2.1.3 *Películas*

Las opiniones que valoran películas tienen un gran impacto en su éxito comercial, debido a la gran influencia que tienen en terceras personas. Esto porque muchas personas generalmente toman en consideración la evaluación y las opiniones, externadas por otras personas, al momento de decidir si consumen o no cierto producto [7].

Determinar con precisión si un comentario es positivo o negativo es una tarea compleja. Además, se debe de considerar que puede llegar a necesitarse clasificar cientos de miles o millones de comentarios acerca de una película. Para esta tarea

se desarrollan clasificadores computacionales que, para su implementación, se requiere emplear técnicas de Inteligencia Artificial. Algunos de los más usados son los algoritmos supervisados, los cuales se entrenan con un conjunto de datos que esta etiquetado manualmente, por expertos, categorizando la polaridad de los comentarios (en positivo o negativo). Entre los algoritmos supervisados más populares, se pueden mencionar: la regresión logística (RL), las máquinas de soporte vectorial (MSV), el método de Naïve Bayes (NB), el de K-vecinos más cercanos (K-VC) y bosque aleatorio (BA).

Considerando lo anterior, el presente trabajo pretende utilizar los algoritmos más utilizados en la metodología dónde se propone llevar a consideraran 2 posibles caminos, el primero que incluye un preprocesamiento y un segundo camino que se considera nuestro “baseline” o punto de partida, que son datos en crudo sin realizar preprocesamiento, ambos lados pasaran por la creación de una bolsa de palabra la cual será analizada en los escenarios de clasificación, los cuales tiene como objetivo determinar cómo se llevaran a cabo los experimentos e indicar la cantidades que se usaran para entrenar a los modelos de aprendizaje automático (k- vecinos más cercanos , regresión logística, Naïve-Bayes, máquinas de soporte vectorial y bosque aleatorio), estos modelos con la configuración establecidas en el escenario de clasificación pasaran por medio de sus fórmulas estadísticas para poder aprender de los conjuntos de datos, y así para poder predecir datos, estas nuevas predicciones serán las que serán evaluadas en la puntuación, esto nos sirve para evaluar una perspectiva sobre el rendimiento de nuestro modelo y nos ofrezcan una visión integral de su eficacia.

Los resultados principales de este estudio se reflejan en la creación de un sistema capaz de clasificar correctamente las categorías de una opinión como positiva o negativa con una exactitud del 97.5%. Además, el sistema tiene una precisión del 97.78% en la predicción de opiniones positivas, lo que significa que hay una probabilidad muy baja de clasificar erróneamente una opinión negativa como positiva. En otras palabras, si el sistema predice una opinión como positiva, es muy probable que sea realmente positiva.

Este equilibrio en la determinación del proceso se refleja en un índice de precisión del 97.78%. Por lo tanto, este sistema se destaca no solo por su alta métrica de precisión, sino también por su exactitud, recall y F1-score, todas superiores al 97%. Esto lo convierte en uno de los mejores sistemas de clasificación disponibles de la actualidad.

## 2.2 Estado del Arte

Para las empresas proveedoras de productos y servicios, así como para las figuras públicas (artistas y políticos, entre otros) es muy importante conocer la opinión de los usuarios y de la ciudadanía en general. En este sentido se han desarrollado un gran número de trabajos encaminados a aportar a la solución de este problema, el cual involucra varias áreas como son: el procesamiento del lenguaje natural, el uso de minería web y el aprendizaje automático. Estos procesos son complejos debido a varios factores, como pueden ser la falta de estructuración de los datos. Por ejemplo, Pérez Abelleira, et al [18] estiman que entre el 80% y el 90% de la información existente en las organizaciones está en forma no estructurada. Para que esta información sea útil, es necesario procesarla y almacenarla con una forma estructurada bien definida, con la finalidad de poder analizar su contenido y de una manera eficiente poder encontrar información de utilidad que de otra forma seguramente sería desconocida [19], [20]. En el caso de opiniones se trata de identificar patrones, tendencias y asociaciones de palabras utilizadas en las opiniones emitidas empleando algoritmos especializados [21]. La actividad de analizar el texto de opiniones o reseñas se trabaja con el objetivo de identificar la emoción detrás de la opinión (a favor, en contra, positiva, negativa, etc.). En este sentido, Shah, et al [9] definen como minería de opiniones al proceso de identificar la emoción a partir de miles de comentarios de manera automática. Ahora, cuando el número de opiniones es muy grande se requiere de herramientas automáticas que realicen esta tarea de manera eficiente. En el ámbito de la clasificación automática de opiniones, se han implementado diversos métodos con el objetivo de incrementar la precisión. Estos métodos combinan técnicas de Aprendizaje Automático (Machine Learning, ML) con enfoques supervisados y no supervisados, aprendizaje profundo, y el uso de diccionarios léxicos [22].

En particular para la industria cinematográfica se emplean procesos de automatización para determinar la polaridad (positiva o negativa) de miles de reseñas individuales. Esto es de gran importancia en la actualidad debido a que puede ayudar a predecir el éxito de una película. Al proceso de automatización de extraer sentimientos expresados en un texto se le conoce como Análisis de Sentimientos (AS, o en inglés SA), mientras que la decisión de la polaridad está relacionada con la Clasificación de Sentimientos (CS, o en inglés SC). Uno de los objetivos de los estudios en el dominio del análisis de opiniones de las reseñas de películas es aumentar la precisión de la clasificación de opiniones en función de los datos de entrada, principalmente en forma de texto [23]. Es importante destacar que

la elección del método o la combinación de métodos a utilizar dependerán en gran medida del tipo de datos disponibles y del contexto específico de la tarea de clasificación de opiniones.

Uno de los primeros trabajos que podemos encontrar dentro del aprendizaje automático es en el uso para el análisis de opiniones en las reseñas de películas de largo metraje en el reportado por Pang, et al [24] para el cual emplearon un conjunto de datos de 1400 reseñas de películas de Hollywood extraídas del IMBD. Este es un popular sitio web de base de datos de películas de los Estados Unidos América y uno de los principales sitios web con contenido acerca de reseñas de películas, de hecho, es el más consultado para reseñas y calificaciones por parte de los usuarios. De acuerdo con los autores, empleando máquinas de vectores de soporte con uni-gramas, lograron alcanzar una precisión del 86%. Posteriormente, Pang et al, en el 2004 [22] mejoraron el método aplicando cortes mínimos y haciendo uso del conjunto de datos de reseñas de películas de Cornell, que es un conjunto de datos de 1400 reseñas de películas de Hollywood de IMDB.

En otro estudio relacionado al aprendizaje automático, es el de Rui, et al. [25] reportan haber usado máquinas de vectores de soporte para alcanzar una clasificación de críticas positivas y negativas de películas con una precisión del 82.9%, superando el 81% obtenido con Naïve Bayes. En este trabajo los autores emplearon el conjunto de datos de críticas de películas de IMDB que consta de 752 críticas negativas y 1301 positivas.

Por otra parte, Martín-Valdivia et al [26] presentaron un modelo para el análisis de sentimientos que combina las características del aprendizaje supervisado y no supervisado, creando el modelo Metaclasificador. Posteriormente, compararon los resultados obtenidos con los diferentes análisis en diferentes idiomas para determinar cuál obtiene mejores resultados. En este estudio realizan la normalización de datos como preprocesamiento, y para el procesamiento hacen uso del léxico SentiWordNet y emplean los métodos de aprendizaje automático: máquinas de soporte vectorial y Naïve Bayes. Además, utilizaron las métricas de exactitud (valores correctamente clasificados), recall (que tan seguros son los positivos clasificados), precisión (verdaderos positivos frente a positivo predichos) y F1 (comparación de rendimiento entre precisión y recall), donde obtuvieron 0.8697, 0.8693, 0.8704, 0.8698, respectivamente para evaluar el método, estos valores se consideran promedio debido a que varios estudios han logrado alcanzar esas métricas. Para su estudio emplearon el conjunto de datos llamado MuchoCine.



Por su parte, So Kim, et al [27] presentaron la propuesta de emplear métodos alternativos de aprendizaje supervisado, incluidos métodos de probabilidad condicional de palabras o polaridad y un método basado en la frecuencia. Al aplicarlos al conjunto de datos de reseñas de películas de IMDB, que incluía 80,000 reseñas de películas, descubrieron que los léxicos que usan probabilidad condicional de palabras muestran una mayor precisión con respecto a los que no la usan.

Además, se han explorado técnicas de aprendizaje profundo, como las redes neuronales convolucionales, para el propósito de mejorar la exactitud. Por ejemplo, en el estudio de Ay Karakus et al [28], se reportó que emplearon varios modelos de aprendizaje profundo para resolver un problema de clasificación de sentimientos binarios. En dicho caso usaron reseñas de películas en idioma turco, del sitio *web www.beyazperde.com*, para entrenar y probar los modelos de aprendizaje profundo. El conjunto de datos completo consta de 44,617 registros, incluidas revisiones positivas y negativas. Se utilizaron únicamente 4000 muestras del conjunto de datos para probar los modelos. Además, los autores emplearon dos modelos principales de aprendizaje profundo: Convolutional Neural Network (CNN) y Long short-term memory (LSTM). Asimismo, para el acoplamiento de las palabras se aplicó el algoritmo word2vec, con un modelo skip-gram. De esta manera, los resultados experimentales demostraron que el uso de la incrustación de palabras con redes neuronales profundas produce mejoras de rendimiento en términos de tiempo de ejecución y precisión.

En otro trabajo relevante para el tema es el realizado por de Duan, et al [29] donde propusieron emplear un modelo de red de tensores neuronales recursivos (RNTN) para identificar oraciones como positivas o negativas mediante el uso de árboles de análisis sintáctico completamente etiquetados. En este trabajo usaron un conjunto de datos de 11,855 reseñas de películas. De esta manera empleando RNTN lograron obtener una exactitud del 80.7% en la predicción del sentimiento en todas las frases y tomar parte de la negación en diferentes sentimientos y alcanzar una mayor precisión que con los modelos anteriores.

## 2.3 Large movie review dataset

El conjunto de datos llamado “Large movie review dataset” fue elaborado por los investigadores Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, y Christopher Potts, en 2011. Este dataset contiene un total de 50

mil opiniones, esto permite el análisis de sentimientos por medio de métodos de aprendizaje tanto supervisados como no supervisado. Esto en principio se debe porque para el aprendizaje supervisado se requiere datos etiquetados, es decir, categorizados en negativos y positivos, y el hecho de tener para ambas categorías 50% y 50% permite un equilibrio a la hora de procesar la información, y para los no supervisados contiene los suficientes datos para poder entrenar a los modelos por medio de sus características es por ello que en varios de los trabajos reportados en la literatura donde se ha utilizado el dataset “Large movie review dataset” generalmente se busca la clasificación a través de diferentes métodos. Por ejemplo, uno de los trabajos pioneros es el de Kumar, et al [30] que por medio de diccionarios léxicos combinados con estadística para la selección de los mejores léxicos busca clasificar las opiniones usando algoritmos como Naive Bayes, Máxima Entropía y máquinas de soporte vectorial. De acuerdo con los autores, la mejor exactitud (89.993%) la obtuvieron usando con el método de Máxima Entropía. En otro interesante trabajo, Timmaraju, et al [31] reportan el uso de redes neuronales recursivas y máquinas de soporte vectorial para la clasificación de opiniones, manteniendo las asociaciones semánticas entre los términos de las palabras, logrando obtener un máximo de exactitud con máquinas de soporte vectorial (MSV) del 86.453%. Por su parte, Ali, et al. [32] reportan que abordaron el problema de la clasificación de sentimientos, usando el dataset “large movie review dataset,” a través de arquitecturas de redes neuronales convolucionales (CNN), con lo que lograron una exactitud del 87.7%, además de investigar por medio de un perceptrón multicapa (MLP) con el cual obtuvieron 86.74% de exactitud en su análisis de nivel de oración y una arquitectura híbrida de CNN con memoria de corto plazo (LSTM) con la que lograron un 89.2% de exactitud. Finalmente, Manek, et. al. [33] en su artículo, proponen el método de selección de características basado en el índice de Gin (índice que mide la impureza) con el clasificador máquina de soporte vectorial para la clasificación de sentimientos para un gran conjunto de datos de reseñas de películas. Los resultados muestran que su método del Índice de Gini tiene un mejor rendimiento de clasificación en términos de menor tasa de error y exactitud, logrando alcanzar un 94.96% de exactitud en su experimentación.

## 2.4 Justificación del proyecto de tesis

La clasificación de polaridad de opiniones es un tema de gran relevancia en particular, para la industria cinematográfica es de relevancia para poder predecir en

función de las opiniones vertidas por usuarios a través de “internet” el éxito de las películas. Sin embargo, este proceso es sumamente complejo de implementar, por lo que diversos grupos de investigación a nivel mundial están buscando formas de mejorar los sistemas de clasificación de polaridad de opiniones. Por lo que en el presente trabajo se propone implementar un sistema, basado en la utilización de diferentes técnicas de inteligencia artificial, que permita clasificar la polaridad de opiniones de películas de largo metraje. Para su implementación se consideran dos escenarios: con y sin etapa de preprocesamiento. Esta distinción consideramos permitirá comparar la precisión de los resultados y la complejidad de la implementación de los métodos. De esta manera se considera que se puede implementar un sistema de clasificación de polaridad de opiniones que sea más simple que algunos de los métodos convencionales y que además sea capaz de hacer la clasificación con una exactitud mayor al 90%.

## 2.5 Preguntas de investigación

- ¿Es indispensable emplear etapas de pre-procesamiento en la implementación de un sistema de clasificación automática de opiniones de largo metraje de manera eficiente?
- ¿Es posible diseñar una nueva metodología, basada en diferentes técnicas de inteligencia artificial, que permita obtener resultados altamente competitivos o superiores a los reportados en la literatura?

## 2.6 Hipótesis

*“Se considera que al implementar un sistema, basado en la utilización de diferentes técnicas de inteligencia artificial, es posible realizar la clasificación automática de sentimientos de las opiniones de películas de largo metraje en dos categorías (positiva y negativa) con una precisión igual o superior a la reportada por métodos similares reportados en la literatura, con el fin de obtener una retroalimentación por parte de los usuarios tomando en cuenta las palabras más relevantes de cada categoría para poder analizar y evaluar diferentes tomas de decisión del éxito o fracaso de las películas.”*

## 2.7 Objetivos

### 2.7.1 *Objetivo General*

- Implementar un sistema basado en técnicas de la inteligencia artificial para identificar automáticamente la categoría positiva/negativa de las opiniones de películas de largo metraje.

### 2.7.2 *Objetivo Específico*

- Desarrollar un sistema basado en técnicas de inteligencia artificial para la identificación de la categoría en opiniones con una exactitud superior al 90%.
- Identificar las técnicas de preprocesamiento y procesamiento que permitan obtener mejores resultados en la identificación del sentimiento en opiniones de películas de largo metraje.

---

## 3 Marco Teórico

---

### 3.1 ¿Qué es la Inteligencia Artificial?

Durante miles de años se ha tratado de entender o asimilar la inteligencia, la forma en las que pensamos y razonamos en corto y largo plazo, este proceso ha sido vital y ha marcado la diferencia en el desarrollo de la sociedad, este conocimiento de aprendizaje se ha combinado con el poder de las computadoras, tratando de simular o lograr la misma comprensión de aprendizaje que un humano, en estas se intenta unir lo mejor de dos entes, la agilidad de la computación y la inteligencia humana. Es así como nace la inteligencia artificial, con el objetivo de lograr que las computadoras realicen tareas que el ser humano realiza de mejor forma, por medio de entidades o sistemas inteligentes [34], por lo tanto, en la búsqueda de poder definir qué es la inteligencia artificial nos retornamos a las palabras de Charniak y Mc Dermott, 1985, el cual nos indica que la inteligencia artificial es:

*“El estudio de las facultades mentales mediante el uso de modelos computacionales”*

Esta afirmación se alinea, con el sistema que se pretende crear y con los modelos de aprendizaje que se hacen uso hoy en día, aun así, existen quienes consideran primordial poner a pruebas las máquinas para evaluar si realmente las respuestas que nos indican se podrían asemejar a un humano, o al aprendizaje de uno, es por ello por lo que en 1950 surge la *Prueba de Turing* con el objetivo de no poder distinguir un humano de una máquina.

La *Prueba de Turing*, propuesta por Alan Turing en 1950, se diseñó para proporcionar una definición operacional y satisfactoria de inteligencia. Esta prueba consiste en diferencias entre entidades inteligentes indiscutibles y seres humanos, el proceso por el cual pasa el computador es una serie de preguntas que deberá responder cuando un humano le pregunte, este computador no podrá ser visible y al mismo tiempo un humano le hará preguntas sin poder ver a su entrevistado, el computador supera la prueba si un evaluador humano no es capaz de distinguir si las respuestas son de una persona o no.

Hoy por hoy, se discute si una computadora que supera la prueba es realmente inteligente, debido a que debería poseer las siguientes capacidades:

- **Procesamiento de lenguaje natural** que permita comunicarse satisfactoriamente.
- **Representación del conocimiento** para almacenar lo que se conoce o siente.
- **Razonamiento automático** para utilizar la información almacenada para responder preguntas y extraer nuevas conclusiones.
- **Aprendizaje automático** para adaptarse a nuevas circunstancias y para detectar y extrapolar patrones [35].

## 3.2 Aprendizaje Automático

Aprendizaje automático (machine learning), es una subdivisión de la inteligencia artificial que permite que los algoritmos detecten patrones recurrentes en un grupo de datos, es considerado como un análisis aumentado que pertenece al subconjunto de la IA. Al encontrar estos patrones los algoritmos aprenden y mejoran los modelos, las predicciones estadísticas o cualquier tarea específica [36]. Los algoritmos de aprendizaje automático son programas informáticos que se dice que aprenden de la experiencia relativa a un conjunto de tareas. Realiza estas tareas con un nivel de precisión que mejora con la experiencia. Es necesario entrenar algoritmos de aprendizaje automático con datos de entrenamiento para obtener el nivel de precisión deseado. Una vez entrenados, estos algoritmos generan conocimientos potentes que pueden utilizarse para predecir resultados futuros. Se concentra en el diseño y desarrollo de algoritmos y técnicas que permiten a las computadoras “aprender” [34].

Se define el término *aprender* como la habilidad de adquirir nuevo conocimiento, desarrollar habilidades para analizar y evaluar problemas a través de métodos y técnicas, así como también por medio de la experiencia propia; siendo un requisito que el resultado del aprendizaje sea entendible para un ser humano. Este proceso se puede expresar en 4 pasos:

- 1) Adquirir nuevo conocimiento.
- 2) Desarrollar habilidades motoras y cognoscitivas por medio de instrucción o práctica.
- 3) Organización de nuevo conocimiento dentro de representaciones generales.
- 4) Descubrimiento de nuevos factores y teorías a partir de la observación y experimentación

Esta forma de procesar información se puede manejar de manera lógica con notación matemática para crear máquinas que razonen y propongan un modelo de solución [37]. De esta forma las aspiraciones por lograr que una computadora interactúe con el humano, como si fuera otro humano, se define como aprendizaje de máquina (machine learning).

En resumen, el aprendizaje automático es un tipo de técnica de inteligencia artificial que se puede utilizar para descubrir automáticamente información útil de conjuntos masivos de datos [38], también conocido como análisis aumentado, se considera un subconjunto de la IA. Está estrechamente relacionado con la estadística computacional y con la estadística en términos de métodos, pero el objetivo subyacente es diferente. Se utilizan técnicas estadísticas descriptivas para hacer inferencias sobre poblaciones (o un conjunto de datos), mientras que los algoritmos de aprendizaje automático encuentran patrones predictivos generales [39].

### 3.3 Tipos de aprendizaje de máquina

El aprendizaje de máquina (machine learning) se clasifica como: 1) supervisado y 2) no supervisado [36].

#### 3.3.1 *Aprendizaje Supervisado*

El aprendizaje supervisado, es cuando entrenamos un algoritmo de Machine Learning dándole las preguntas (características) y las respuestas (etiquetas). Así en un futuro el algoritmo pueda hacer una predicción conociendo las características [40]. Constituye un algoritmo de aprendizaje basado en ejemplos donde el nuevo conocimiento es inducido a partir de una serie de ejemplos y contraejemplos. es decir, este tipo de aprendizaje se basa en un conjunto de datos de entrenamiento etiquetados, lo que significa que cada muestra de datos del conjunto de entrenamiento tiene una etiqueta que indica la respuesta correcta que previamente un experto del tema categorizo manualmente cada característica y produce una función que establece una correspondencia entre las entradas y las salidas deseadas del sistema. Un ejemplo de este tipo de algoritmo es el problema de clasificación y regresión, el primero se espera que el algoritmo nos indique a que grupo pertenece la característica mientras que el segundo el algoritmo nos regresa un valor numérico, un ejemplo de cada uno es el siguiente: para clasificar el sistema de aprendizaje trata de etiquetar (clasificar) una serie de datos, a partir del

conocimiento que aprendió del conjunto de entrenamiento donde se le indicaba que tipo de dato corresponde a cada categoría, mientras que el otro algoritmo tiene el precio de diferentes casas, pequeñas, grandes, en el campo, en la ciudad, etc. y por medio de un gráfico de dispersión, puede predecir el precio correcto de una casa en consulta [34], [40].

### 3.3.2 *Aprendizaje no supervisado*

El aprendizaje no supervisado, es el que no tiene información sobre las categorías de ejemplo, ni conjunto de entrenamiento que lo ayuden a distinguir clases o categorías. Constituye un tipo de aprendizaje donde el sistema de aprendizaje analiza una serie de entidades y determina que algunas entidades tienen características comunes, por lo que pueden ser agrupadas formando un concepto y a partir de estos conceptos aprende nuevos. Un ejemplo de este tipo de algoritmo es el problema de clasificación que desempeñan las redes neuronales artificiales [34].

## 3.4 Minería de datos

La minería de datos se refiere a la extracción de información valiosa que puede ser representada en una forma de conocimiento estructurado y comprensible, como gráficos y patrones. Las técnicas de minería de datos persiguen el descubrimiento automático del conocimiento contenido en la información almacenada de modo ordenado en grandes bases de datos. Estas técnicas tienen como objetivo descubrir patrones, perfiles y tendencias a través del análisis de los datos utilizando tecnologías de reconocimiento de patrones, redes neuronales, lógica difusa, algoritmos genéticos y otras técnicas avanzadas de análisis de datos [41].

Esta medida ha provocado que los negocios mundiales se hagan más competitivos, dando a los datos cada vez más importancia y convirtiendo la información en algo vital y estratégica para la toma de decisiones. En tal sentido, las empresas han venido evolucionando y han querido agregarle valor a la gran cantidad de información que tienen almacenada en sus bases de datos. Para ello, se han interesado en automatizar los procesos y poder así descubrir información valiosa, que de otra manera seguiría siendo subutilizada o simplemente desperdiciada. Con la gran velocidad a la cual ha venido evolucionando la tecnología, las empresas de



hoy disponen de herramientas de software y hardware cada vez más sofisticadas que posibilitan el almacenamiento de grandes cantidades de información y el análisis de esta. El avance tecnológico, sumado a la aparición de mercados cada vez más competidos, sugiere a las empresas el mejoramiento continuo de sus esquemas de administración y toma de decisiones, explotando una de las más grandes fuentes de competitividad; la información [42]. Es por ello por lo que el principal enfoque de las compañías es desarrollar sistemas de información, con el fin de dar soporte a los procesos básicos de la organización (ventas, producción, personal...). Una vez satisfecha la necesidad de tener un soporte informático para los procesos básicos de la organización (sistemas de información para la gestión), las organizaciones exigen nuevas prestaciones de los sistemas de información (sistemas de información para la toma de decisiones) para aprovechar la información de las opiniones que generan sus productos, buscando encontrar una plusvalía o ventaja competitiva, generando la extracción de las de opiniones y dando lugar a la minería de opiniones [43].

### *3.4.1 Minería de opiniones*

La minería de opiniones, también conocida como análisis de sentimientos, se ocupa de la extracción, identificación y caracterización de las opiniones expresadas en un texto. Esta rama de la minería de datos ha ganado importancia debido al crecimiento exponencial de los datos generados por los usuarios en diversas plataformas en línea, como las redes sociales y los sitios de reseñas [43], [44].

La evolución de la extracción de información ha sido un proceso fascinante que se ha desarrollado a lo largo de más de tres décadas. Esta disciplina, que comenzó con la simple extracción de información de textos, ha evolucionado hasta convertirse en lo que hoy conocemos como minería de opiniones. En la década de 1990, la Message Understanding Conference (MUC) declaró que la "Extracción de Información" (IE) era una técnica automática para desarrollar una representación bien definida de la información recopilada a partir de documentos de texto. La extracción de información de textos, artículos de noticias y redes sociales implica principalmente el uso del procesamiento del lenguaje natural (PLN) [45].

Con el avance de la Web y la disponibilidad de Internet, se ha prestado considerable atención a los documentos Web, ya que sirven como un medio innovador para divulgar experiencias individuales en forma de reseñas, opiniones, blogs, comentarios o retroalimentaciones. Estos son en sí mismo "medios generados por

el consumidor". Las redes sociales se han convertido en una necesidad social de comunicación de la sociedad actual [46]. Hasta hace unos cuantos años, la gente tendía a vender productos mediante la publicación de anuncios y se recababan las opiniones realizando: encuestas, grupos focales, consultores, tomando en cuenta las opiniones de amigos y familiares, estas herramientas tenían la desventaja que usualmente la retroalimentación podía llevar cierto tiempo y a los usuarios no les llamaba la atención responder o participar en las retroalimentaciones, considerando más fácil calificar un producto o dejar una reseña [45]. Iniciando la extracción por medio de técnica como la IE que realizaba relativamente fuertes análisis de subjetividad o análisis de sentimiento utilizando diversas técnicas estadísticas y de PNL [44], creando un efecto en los usuarios de la Web que buscan reseñas en línea para tomar decisiones sobre ciertos productos de su interés debido a la fácil disponibilidad de crear y ver reseñas en línea, siendo ahora la fuente principal de mejora de la calidad de los servicios por medio de la opinión de los usuarios.

La creciente popularidad de los sitios de compras en línea, las redes sociales y otros sitios de interacción de reseñas ha hecho que millones de usuarios publiquen opiniones en la World Wide Web (WWW). Cada vez que una persona compra un producto, puede recopilar información directamente en la Web. Durante las elecciones políticas, las decisiones de votación se pueden tomar de manera similar. Las investigaciones de mercado se pueden realizar mediante la extracción de información en Internet para recopilar conocimientos e interés de los clientes. Todos estos factores apuntan a centrarse de manera más inclusiva en la minería de opiniones. Esto ha llevado a generar interés en las tecnologías y en la investigación para extraer, clasificar y resumir automáticamente opiniones a partir de documentos web y, por lo tanto, ha llevado al desarrollo de sistemas que extraen opiniones [44], [47].

Muchos investigadores han aplicado algoritmos de aprendizaje automático supervisados/no supervisados bien conocidos para clasificar y agrupar datos con el fin de determinar el sentir de las personas. Es ahora cuando surge la necesidad de desarrollar un sistema que pueda proporcionar a los usuarios este tipo de reseñas de manera resumidas, entonces la sociedad puede beneficiarse de varias maneras:

- Reduce los esfuerzos y el tiempo que las personas dedican a buscar reseñas confiables.
- Las personas pueden ver las opiniones según las características que les interesan.

A partir de este punto el análisis de opiniones viene de la búsqueda de la información que permite a las organizaciones identificar que características de sus productos y servicios son las que más les gusta y no les gusta. Considerando una visión más amplia, la minería de opiniones y el análisis de sentimientos van de la mano y representan el mismo campo de estudio [47].

### 3.5 Dicionarios léxicos

Los diccionarios léxicos son bases de datos léxicas y con etiquetas que permiten identificar las partes de la oración (sujeto, verbo, sustantivo), para posteriormente poder hacer la clasificación del sentimiento asociado, su uso se asocia principalmente en la clasificación de opiniones especialmente cuando se dispone de un texto grande y diverso.

Existen diferentes enfoques de la clasificación de opiniones, entre las que se encuentran el uso de léxicos de opiniones (se refiere a un conjunto de palabras y frases que se utilizan para expresar opiniones, sentimientos o juicios de valor). Existen léxicos (conjuntos de palabras que una persona usa cotidianamente) ya contruidos, pero también se pueden crear específicamente para que nos proporcionen una puntuación en la orientación de opiniones [48]. Este enfoque tiene la ventaja de evitar el arduo paso de etiquetar (marca que se usa para identificar, clasificar o valorar un objeto) los léxicos de los datos de entrenamiento (valores asociados a una etiqueta). Sin embargo, estas técnicas se basan en recursos léxicos externos (es decir, en el campo de la lexicografía digital, un recurso léxico es un recurso lingüístico que consta de datos sobre los lexemas [un conjunto de formas tomadas por una sola] del léxico [catálogo de palabras de un idioma] de uno o más idiomas) que se ocupan de asignar palabras a una puntuación de opinión categórica (p.e. positiva, negativa, neutral, entre otros) o numérica, que utiliza el algoritmo para obtener la opinión general que transmite el texto [49]. Por ejemplo, palabras como “excelente”, “malo”, “mejor”, “peor” son parte de un léxico de opiniones porque expresan juicios de valor. En resumen, un léxico de opiniones es una herramienta que nos permite entender mejor cómo las personas expresan sus opiniones y sentimientos a través del lenguaje. Por lo tanto, la eficacia de todo el enfoque depende en gran medida del recurso léxico empleado. Actualmente, existen varios léxicos de sentimientos genéricos entre los que destacan: SentiWordNet [50], MPQA Subjectivity Lexicon [51], Linguistic Inquiry and Word

Count (LIWC) [52], SentiStrength [53], SentiWords [54], Affective standards for English Word (ANEW)[55] y General Inquirer [56]. Estos se han utilizado reciente y popularmente en tareas de análisis de sentimientos. Por otra parte, los léxicos se pueden crear manual o automáticamente, por ejemplo, usando palabras semilla para expandir la lista de palabras [57], [58], [59]. Asimismo, la mayoría de los estudios utilizan el concepto de PMI (Pointwise Mutual Information) para crear un léxico a partir de un conjunto de palabras semilla.

### 3.6 Análisis de Sentimientos

Tradicionalmente la computación afectiva ha estudiado el problema del análisis de sentimientos con el fin de analizar la opinión y el sentimiento de las personas hacia entidades, como productos, servicios, entre otros [60]. Este dominio se ha estudiado, en gran medida, debido a la gran cantidad de aplicaciones del mundo real en las que es importante descubrir la opinión de las personas. Considerando que hay una gran cantidad de datos generados en diferentes dominios, que expresan información específica de una *entidad* (producto), y que son necesarios de analizar para conocer el impacto del producto en los usuarios potenciales. Estos comentarios son de gran importancia para los procesos de aprendizaje de una persona, debido a que un sentimiento (negativa o positiva) de una entidad puede influir en el estado emocional de los usuarios, y decidir si adquirir o no el servicio que se ofrece, dando así que una entidad puede promover estados afectivos durante el proceso de aprendizaje de las personas, jugando un papel clave en la motivación, y la forma de promover entidades, que provoca que se pueda analizar en otros contextos los diferentes sentimientos de una entidad [47].

Ahora bien, el análisis de sentimientos es el encargado de clasificar un documento de diferentes contextos de aplicación (productos, películas, deportes, culturales, etc.) en función de la orientación semántica de la opinión que expresa el autor, mediante enfoques semánticos (diccionarios de términos), aprendizaje de máquina (por ejemplo: máquinas de soporte vectorial, Naïve Bayes, aprendizaje profundo, etc.). En este contexto, la determinación de polaridad de una opinión es valorada como positiva, neutra o negativa acerca de un producto, servicio, organización, persona o sobre cualquier tipo de ente sobre el que es posible expresar de manera escrita una opinión. Un enfoque típico para el análisis de sentimientos de los

documentos es utilizar las frecuencias de palabras positivas y negativas, para determinar si un documento es predominantemente positivo o negativo

En el ámbito de la computación según Liu B. [61] las opiniones pueden describirse como una "quíntupla" formada por cinco conceptos:

- ✓ el objeto sobre el que se basa la opinión,
- ✓ la característica del objeto a la que se hace referencia,
- ✓ el sentimiento sobre esta característica,
- ✓ el opinador
- ✓ el momento en el tiempo en el que se hace la opinión.

De conseguir estas "quíntuplas" en el análisis de un texto habríamos conseguido pasar de una información sin estructura inicial a una estructurada por la cual nuestro sistema PLN podría actuar. Ejemplos relacionados con las opiniones de películas se muestran en la Figura 1.

```
{Robocop, entretenimiento, positivo, ademir1, 11-ene-2005}  
{Avatar, 3D, positivo, object_52, 23-mar-2015}  
{Birdman, humor, negativo, raul53#, 21-abr-2020}
```

Figura 1. Ejemplo de una quintupla en un análisis de Alcalde, 2021.

Los sistemas PLN y concretamente los dedicados a análisis de sentimiento tienen que considerar también el tipo de texto a analizar. Según un caso u otro el sistema tendrá que ajustar su procesado [47], [62]. Básicamente, se tienen tres tipos posibles de orígenes de información a analizar:

- i. **Documentos.** La tarea en este nivel es clasificar si un documento de opinión completo expresa un sentimiento positivo o negativo. Por ejemplo, dada una reseña de un producto, el sistema determina si la reseña expresa una opinión general positiva o negativa sobre el producto. Esta tarea se conoce comúnmente como clasificación de sentimiento/nivel de documento. Este nivel de análisis supone que cada documento expresa opiniones sobre una sola entidad (por ejemplo, un solo producto), es un caso en el cual se ha de realizar todo un proceso de desglose y clasificación de conceptos (en inglés document-level sentiment classification). Por lo tanto, no es aplicable a documentos que evalúan o comparan múltiples entidades.

- ii. **Frases.** Nivel de oración: la tarea en este nivel va a las oraciones y determina si cada oración expresó una opinión positiva, negativa o neutral. Neutral generalmente significa no tener opinión. Este nivel de análisis está estrechamente relacionado con la clasificación de la subjetividad, que distingue oraciones objetivas, que expresan información fáctica, y las oraciones subjetivas, que expresan puntos de vista y opiniones subjetivos. Sin embargo, se debe tener en cuenta que la subjetividad no es equivalente al sentimiento, ya que muchas oraciones objetivas pueden implicar opiniones, por ejemplo, "*Compramos el auto el mes pasado y el limpia parabrisas se cayó*". Los investigadores también han analizado las cláusulas, pero el nivel de las cláusulas aún no es suficiente (por ejemplo, "A Apple le está yendo muy bien en esta pésima economía"). Este tipo de análisis se centraría sobre todo en los aspectos de subjetividad y objetividad por las cuales se puede expresar una opinión.
  
- iii. **Entidades/aspecto:** Tanto el análisis a nivel de documento como el de oración no descubren exactamente qué le gustó y qué no le gustó a la gente. El nivel de aspecto realiza un análisis más detallado. El nivel de aspecto anteriormente se llamaba nivel de característica y hace una extracción y resumen de opiniones basado en características, en lugar de observar construcciones del lenguaje (documentos, párrafos, oraciones, cláusulas o frases). Por lo que se dice que el nivel de aspecto observa directamente la opinión misma.

Darnos cuenta de la importancia de los objetivos de opinión también nos ayuda a comprender mejor el problema del análisis de sentimientos. Por ejemplo, aunque la frase "Aunque el servicio no es tan bueno, todavía me encanta este restaurante". tiene claramente un tono positivo, no podemos decir que esta frase sea enteramente positiva. De hecho, la frase es positiva sobre el restaurante (enfaticada), pero negativa sobre su servicio (no enfaticada). En muchas aplicaciones, los objetivos de opinión se describen por entidades y/o sus diferentes aspectos. Por lo tanto, el objetivo de este nivel de análisis es descubrir sentimientos sobre las entidades y/o sus aspectos. Por ejemplo, la frase "La calidad de la llamada del iPhone es buena, pero la duración de la batería es corta" evalúa dos aspectos: la calidad de la llamada y la duración de la batería de un iPhone (entidad). El sentimiento sobre la calidad de las llamadas del iPhone es positivo, pero el sentimiento sobre la duración de la

batería es negativo. La calidad de las llamadas y la duración de la batería del iPhone son los objetivos de opinión. A partir de este nivel de análisis se puede producir un resumen estructurado de opiniones sobre entidades y sus aspectos, que convierte texto no estructurado en datos estructurados y puede utilizarse para todo tipo de análisis cualitativos y cuantitativos [43].

### 3.7 Orientación semántica

La orientación semántica se refiere a asignar un significado a un significante (palabra) que ha sido establecido arbitrariamente para denominar un elemento o concepto. Llamadas también palabras de sentimiento, o de opinión, o palabras que contienen opinión, estas pueden referirse tanto a palabras individuales como a frases. Las palabras de opinión positiva se usan para expresar estados deseados mientras que las palabras de sentimiento negativo se usan para expresar cualidades o estados no deseados. Algunos ejemplos de palabras de sentimiento positivo son: hermoso, maravilloso, bueno, y asombroso. Ejemplos de palabras de opinión negativas son: malo, pobre y terrible [44], [62].

### 3.8 Modelos de aprendizaje automático.

El conocimiento extraído en forma de relaciones, patrones o reglas inferidos de los datos y (previamente) desconocidos, o en forma de una descripción más concisa (es decir, su resumen) se denomina modelo [63].

#### 3.8.1 *Naïve Bayes*

El algoritmo Naïve Bayes (NB) es un clasificador probabilístico basado en el teorema de Bayes con una suposición de independencia entre los predictores pertenece a los métodos de aprendizaje supervisado, donde por medio de varios datos de entrenamiento los cuales poseen una clasificación predecir de una colección de datos sin etiquetar a que categoría pertenecen [64], [65].

El clasificador Naïve Bayes es un método basado en el teorema de Bayes. Es uno de los clasificadores más usados debido a la independencia de cada condición,

además de que el modelo se puede implementar con grandes volúmenes de datos [66]. La base del teorema utiliza la fórmula de Bayes, definida por la Ecuación (1):

$$P(C|X) = \frac{P(C)*P(X|C)}{P(X)} \quad (1)$$

donde  $X$  significa los atributos,  $C$  las clases,  $P(C|X)$  la probabilidad de que sea  $C$  dado  $X$ ,  $P(X|C)$  probabilidad de que sea  $X$  dado  $C$ ,  $P(C)$  probabilidad de que sea  $C$  y finalmente  $P(X)$  probabilidad de que sea  $X$  [65].

Este proceso se puede entender mejor con un ejemplo: se imagina que se tiene un conjunto de correos electrónicos y quieres clasificarlos en “spam” o “no spam”, en la Figura 2 se puede mostrar un conjunto de datos, en el cual tenemos 2 categorías, una de “spam”, la cual contiene 4 datos etiquetados y una de “no spam” que contiene 7 elementos también etiquetados.

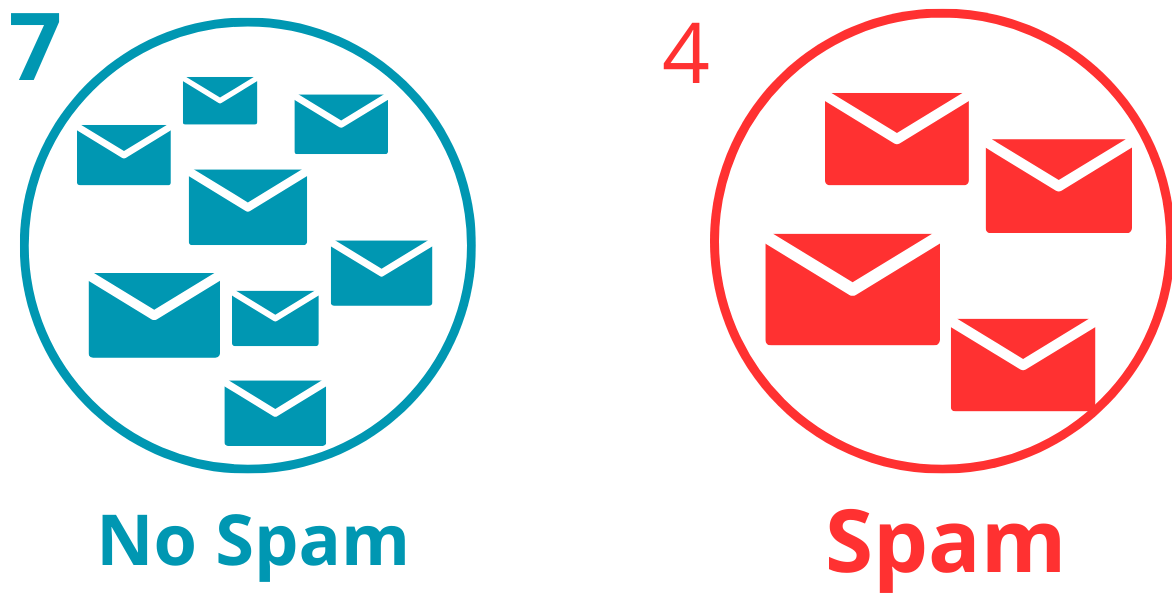


Figura 2. Ejemplo de conjunto de datos para Naive Bayes

Ahora que tenemos nuestro conjunto de datos, debemos analizar las palabras que contiene en este caso hipotético diremos que estos correos solo tienen 3 palabras; “oferta”, “amigo” y “gratis” como se muestra en la Tabla 1, estas palabras nos ayudaran a determinar la probabilidad de que si un elemento tiene estas palabras podamos saber que categoría es, con base a las veces que aparece esa palabra en una categoría.



Tabla 1. Ejemplo de datos con etiqueta para Naive Bayes

| Mensaje | Contenido | Etiqueta | Contenido     | Etiqueta |
|---------|-----------|----------|---------------|----------|
| 1       | Amigo     | No spam  | Oferta Gratis | Spam     |
| 2       | Amigo     | No spam  | Gratis Oferta | Spam     |
| 3       | Amigo     | No spam  | Gratis Gratis | Spam     |
| 4       | Amigo     | No spam  | Oferta        | Spam     |
| 5       | Oferta    | No spam  |               |          |
| 6       | Oferta    | No spam  |               |          |
| 7       | Gratis    | No spam  |               |          |

Tomando en cuenta la ecuación 1, podemos ver que requerimos de 3 elementos:

1.  $P(X|C)$  probabilidad de que sea  $X$  dado  $C$ .
2.  $P(C)$  probabilidad de que sea  $C$ .
3.  $P(X)$  probabilidad de que sea  $X$ .

Para poder entender mejor estas probabilidades, iniciamos con el primer elemento  $P(X|C)$  la probabilidad de que sea  $C$ , en esta probabilidad requerimos de 2 elementos, una clase llamada  $C$  y el atributo, simplificándolo a nuestro ejemplo,  $C$  es la categoría, es decir, "Spam" (S) y "No spam" (N), los marcaremos con un color como se mostró en la Figura 2, el S tendrá color rojo y el N pasara a ser azul, entonces al remplazar estas variables por las categorías nos quedaría  $P(N|X)$  la probabilidad de que sea "N" es decir "No spam", dado el atributo, este atributo son las palabras que buscamos "oferta", "amigo" y "gratis", entonces remplazamos por cada una de estas palabras para cada categoría es decir debemos tener 6 probabilidades, 3 probabilidades de que la palabra sea "No spam" y otras 3 para "Spam", en la Figura 3 se puede ver estas 6 probabilidades por categoría, pero ahora requerimos de un elemento más, sacar la probabilidad de esta condición, para esto debemos saber cuántas veces se repite la palabra en el conjunto de datos por categoría, esto nos recrea un histograma, que es un gráfico que representa la frecuencia, en este caso la frecuencia de palabras por categoría, recordando un poco cómo sacar esta condición, se realiza a partir de  $\frac{\text{numero de eventos donde sucede}}{\text{total de eventos de la categoria}}$ , dando el resultado de la Figura 3.

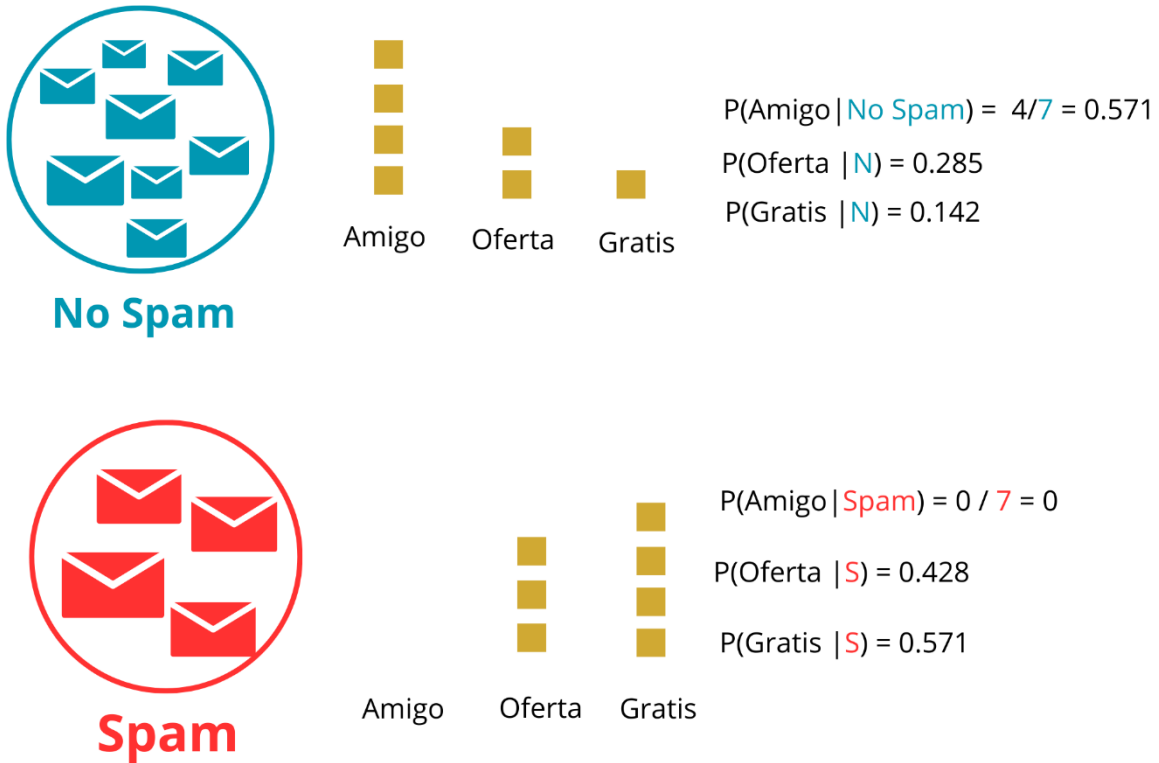


Figura 3. Ejemplo con histograma y probabilidades

Después, nos hace falta conocer la  $P(C)$  la probabilidad de que sea una categoría u otra, esto se hace obteniendo la probabilidad de cada clase, es decir sacar la probabilidad de que sea “spam” y “no spam”, recordando probabilidad, la fórmula para obtener este resultado se realiza en la siguiente ecuación:  $\frac{\text{numero de evento donde sucede}}{\text{total de los eventos}}$ , como se puede observar en la Figura 4, se calculan estas probabilidades para poder sacar el 2 elemento.

7



No Spam

$$P(N) = \frac{7}{7+4} = 0.636$$

4



Spam

$$P(S) = 0.363$$

Figura 4. Ejemplo de probabilidad para cada conjunto

Para calcular el 3 elemento se requiere aplicar una suma de probabilidades, debido a que se encuentra en 2 categorías, debemos sacar la probabilidad del atributo (la palabra) y sumar la probabilidad condicional de ambas categorías dado que sean spam y no spam, como se aprecia en la Figura 5, para tener nuestros 3 elementos y poder sustituir en la Ecuación 1, esto es aplicando el teorema, pero para la clasificación se hace uso de otro método.

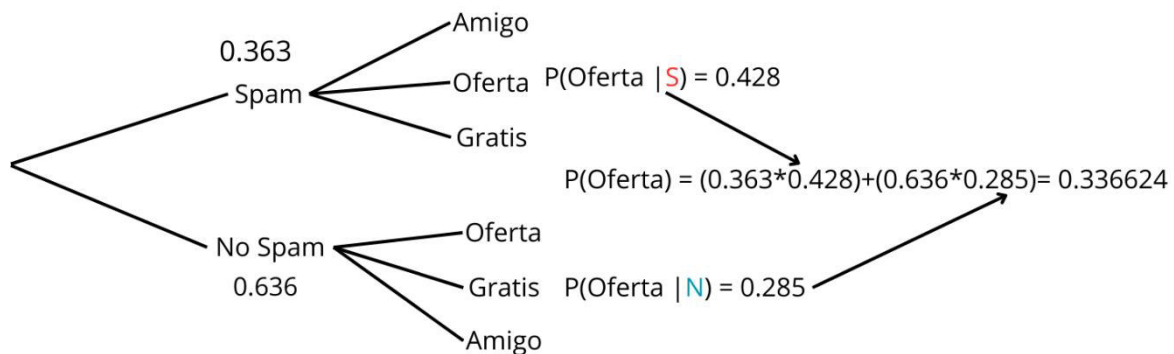


Figura 5. Probabilidad de X en el ejemplo.


Finalmente reemplazamos los elementos, en la Figura 6 se puede ver un ejemplo de este proceso con la palabra oferta, en esta figura se representa que es probable que la palabra oferta sea clasificada como “no spam”.

$$P(S|Oferta) = \frac{0.363 * 0.428}{0.336624} = 0.46153$$

$$P(N|Oferta) = \frac{0.636 * 0.285}{0.336624} = 0.53846 \quad \checkmark$$

Figura 6. Sustitución de los valores en ecuación

Para la clasificación de nuevos correos electrónicos, si se recibe un nuevo correo electrónico y se quiere clasificar como “spam” o “no spam”, se puede aplicar el teorema de Bayes. Para cada clase (es decir, “spam” y “no spam”), se multiplica la probabilidad de esa clase por las probabilidades condicionales de cada palabra en el correo electrónico dado esa clase. La clase con el valor más alto es la clase asignada al correo electrónico, como se muestra en el siguiente Figura 7. En este caso se clasificaría como “no spam” porque la categoría con mayor valor es la de “no spam” pero aquí ocurre un error y es debido a que la probabilidad de que sea amigo en spam tiene valor de 0 esto a la larga puede provocar una mala clasificación y para evitar este error se agrega un valor a cada histograma para poder tener mayor certeza Figura 8.



$$P(N) * P(\text{Amigo} | N) * P(\text{Oferta} | N) = 0.103$$

$$P(S) * P(\text{Amigo} | S) * P(\text{Oferta} | S) = 0$$

Figura 7. Ejemplo para determinar categoría de un nuevo mensaje

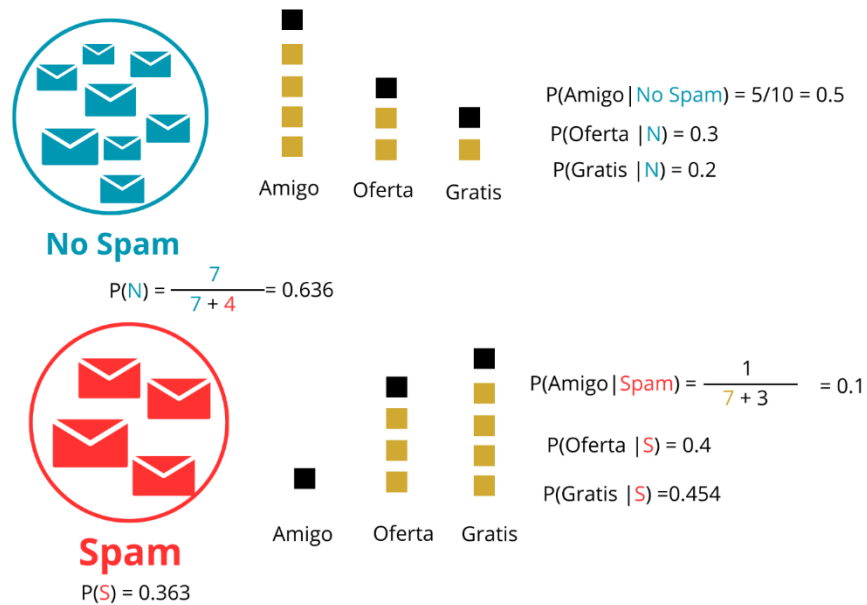


Figura 8. Histograma con adición de un elemento para mejor clasificación

Este proceso solo cambia para la probabilidad de las palabras nuestra probabilidad de “spam” y “no spam” se queda igual, ahora el valor del nuevo correo es Figura 9. Por lo que, el nuevo correo si pertenece a la categoría de “no spam”.

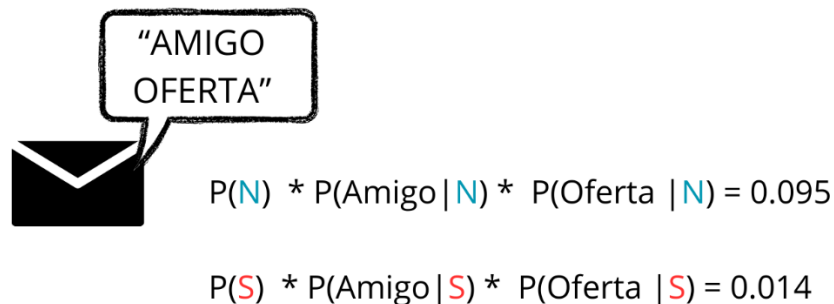


Figura 9. Ejemplo con histograma cambiado para determinar categoría de nuevo mensaje

### 3.8.2 Regresión Logística

La Regresión Logística (RL) es un algoritmo de análisis y utilizado para la clasificación, es utilizado cuando el objetivo es describir la relación entre una variable respuesta categórica, y un conjunto de variables explicativas que pueden ser categóricas o cuantitativas, para poder estimar que cierto evento ocurra teniendo en cuenta las demás variables [67], [68].

Su modelado está en función de  $\pi(x)$ , donde podemos obtener resultados de 0 a 1, siendo 0 la representación de un 0% probable de que un evento ocurra en el estudio y 1 significa que es 100% probable de que suceda. Esto se puede expresar por medio de la Ecuación (2) [67], [68]:

$$P(y = 1|X) = \pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}} \quad (2)$$

donde

$$g(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (3)$$

donde  $\beta_0$  es la constante del modelo o termino independiente,  $p$  el número de covariables,  $\beta_i$  los coeficientes de las covariables, y  $x_i$  las covariables que forman parte del modelo.

Si alguna de las variables independientes es una variable discreta con  $k$  niveles se debe incluir en el modelo como un conjunto de  $k-1$  “variables de diseño” o “variables dummy”. El cociente de las probabilidades correspondientes a los dos niveles de la variable respuesta se denomina odds y su expresión está definida por la Ecuación. (4):

$$\frac{P(y=1|X)}{1-P(y=1|X)} = e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p} \quad (4)$$

Al aplicar el logaritmo natural, se obtiene el modelo de regresión logística, expresado por la Ecuación (5) [67]:

$$\ln \left[ \frac{P(y=1|X)}{1-P(y=1|X)} \right] = \ln(e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}) \Rightarrow \ln \left[ \frac{P(y=1|X)}{1-P(y=1|X)} \right] = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (5)$$

Para ilustrar, de una manera más efectiva el método de regresión logística, consideremos el siguiente escenario: Supongamos que un educador desea predecir la probabilidad de que un estudiante apruebe un examen basándose en la cantidad de horas dedicadas al estudio. En este contexto, “suspender” se consideraría como un resultado de «fracaso», mientras que “aprobar” se consideraría como un resultado de «éxito». La variable independiente en este caso sería el número de horas que un estudiante ha dedicado al estudio.

El primer paso sería la recopilación de datos. Se necesitaría recoger información sobre los estudiantes, incluyendo si aprobaron o suspendieron el examen y cuántas horas dedicaron al estudio para el examen. Posteriormente, se utilizarían estos datos para ajustar un modelo de regresión logística. Esto proporcionaría una ecuación que se podría utilizar para predecir la probabilidad de aprobar el examen en función de las horas de estudio. La ecuación de la regresión logística podría verse así:

$$\ln \left[ \frac{p(X)}{1 - p(X)} \right] = \beta_0 + \beta_1 x$$

Donde:

$p(X)$  es la probabilidad de aprobar el examen.

$\beta_0 + \beta_1 x$  son los coeficientes de la regresión que se aprenden a partir de los datos.

$x$  es el número de horas de estudio.

Por ejemplo, si  $\beta_0 = -2$  y  $\beta_1 = 0.5$ , entonces la ecuación sería:

$$\ln \left[ \frac{p(X)}{1 - p(X)} \right] = -2 + 0.5x$$

Esto significa que, por cada hora adicional de estudio, el logaritmo de las probabilidades de aprobar el examen aumenta en 0.5.

Finalmente, con la ecuación obtenida, se podría predecir la probabilidad de que un estudiante apruebe el examen en función de cuántas horas ha estudiado. Por ejemplo, si un estudiante ha estudiado 4 horas, se podría sustituir  $x = 4$  en la ecuación para obtener la probabilidad de aprobar el examen, en el (Figura 10) se puede observar el resultado de la regresión considerando diferente número de horas de estudio.



Figura 10. Ejemplo de gráfico de regresión logística

### 3.8.3 Máquinas de soporte vectorial

Las máquinas de soporte vectorial (MSV) es un algoritmo de aprendizaje supervisado que se puede utilizar tanto para fines de clasificación como de regresión. La idea de una máquina de soporte vectorial es generar un hiperplano que permita separar una clase de otra maximizando la distancia entre los puntos de diferentes clases y una función separadora [64].

Una MSV no lineal usa varias funciones *kernels* para estimar el margen, que es la distancia desde el punto más próximo de una clase a la función separadora. Los kernels pueden estar descritos por diferentes funciones, como por ejemplo una recta, un polinomio, una función de base radial o una sigmoidea. El principal objetivo de estas funciones es maximizar el margen entre los hiperplanos.

La Ecuación (6) permite aproximar  $y_t$  a través de la función:

$$y_t = b + \sum_{d=i}^D w_d k(x_T, x_d) \quad (6)$$

donde  $x_T$  son los regresores o las variables que representan el problema,  $D$  es el número de ejemplos representativos con los que cuenta la base de datos,  $b$  es una constante y  $w_d$  son los factores de ponderación de la función de núcleo  $k(x_T, x_d)$ . De esta forma, una MSV es la combinación lineal de la imagen de  $x_T$  en un espacio de características no lineal definido por los puntos  $x_d$  y la función de transformación no lineal  $k(x_T, x_d)$  [69].

Para entenderlo de manera intuitiva, consideremos un ejemplo sencillo. Suponga que se está conformando un equipo de ciclismo y se deben seleccionar los



corredores, los cuales pueden pertenecer a una de dos categorías: escalador (especialista en la montaña) o embalador (especialista en las llegadas con terreno plano). La asignación de la categoría a la que pertenece cada ciclista puede tomarse con base en diferentes criterios o características. Por ejemplo, el peso y la potencia que desarrolla el ciclista en cada pedalazo. En dos dimensiones, se puede trazar una línea que divida a estos dos grupos, por ejemplo, en la Figura 11 se puede ver con mayor claridad este caso hipotético de manera visual.

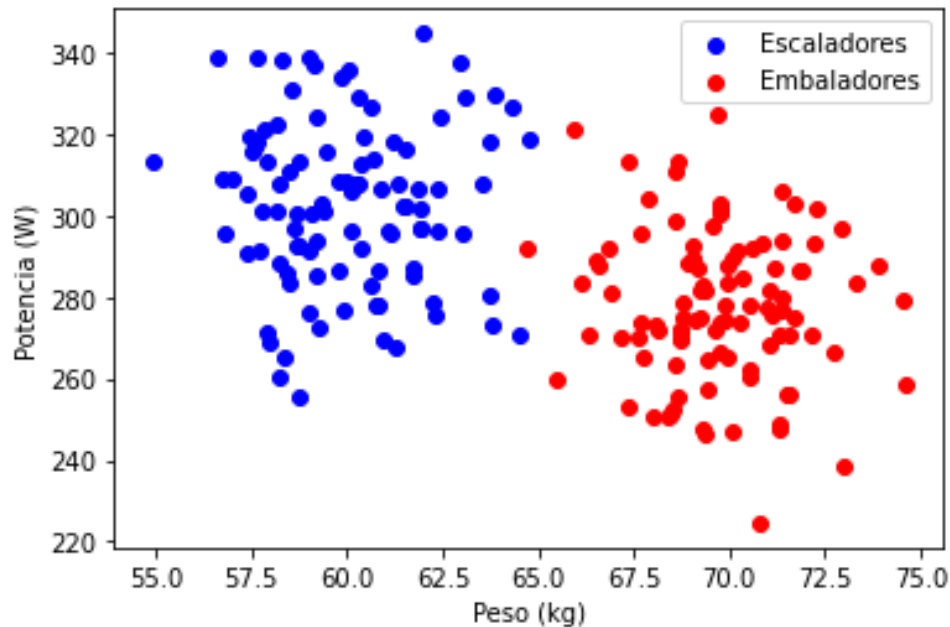


Figura 11. Ejemplo de un gráfico por categorías

Si además del peso y la potencia incluimos la capacidad pulmonar, entonces tendremos una distribución de puntos en tres dimensiones (es decir, tres características), y en este caso tendremos un plano que separa una categoría de otra. En todos los casos, idealmente tendremos una frontera de decisión que permite determinar la categoría a la que pertenece cada ciclista. Esta frontera de decisión es lo que las SVM buscan maximizar, entendiendo el margen como la distancia más corta entre la frontera de decisión y cualquiera de las muestras.

#### 3.8.4 Método de K-Vecinos más cercanos

El método de K-vecinos más cercanos (en inglés K-nearest neighbour, KNN, o K-VC) es un algoritmo utilizado para la clasificación, reconocimiento de patrones y consistencias de los datos, es un algoritmo de aprendizaje supervisado. Este

algoritmo requiere de un conjunto de datos de entrenamiento para poder generar un modelo de clasificación. En este algoritmo, trazamos cada observación de datos como un punto en un espacio n-dimensional (donde n es el número de características que tiene), los vecinos de estos puntos se derivan de un objeto conocido en el espacio dimensional intentando agrupar los objetos similares lo más cerca posible entre sí. [70], [71]. K-NN agrupa en un plano dimensional los puntos de todos los casos de clasificación etiquetados y a partir de ahí, ayuda a clasificar nuevos casos con la medida de similitud. Para calcular la distancia entre puntos se usa la formula euclidiana de K-NN Ecuación (7) [72]

$$d(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2} \quad (7)$$

Un ejemplo de esto lo podemos observar en la Figura 12, en esta contiene dos grupos de círculos, que previamente se conocía su categoría, y agregamos un punto que será conocido como k, este k contará los vecinos que tiene para identificar que categoría pertenece él, en este caso se observa que tenemos 4 círculos amarillos, mientras solo tenemos 2 azules, esta votación determina que el pertenece a la categoría de color amarillo, entonces ahora k pasa a ser amarillo, esto proceso se repite con los datos desconocidos y así predecimos los datos.

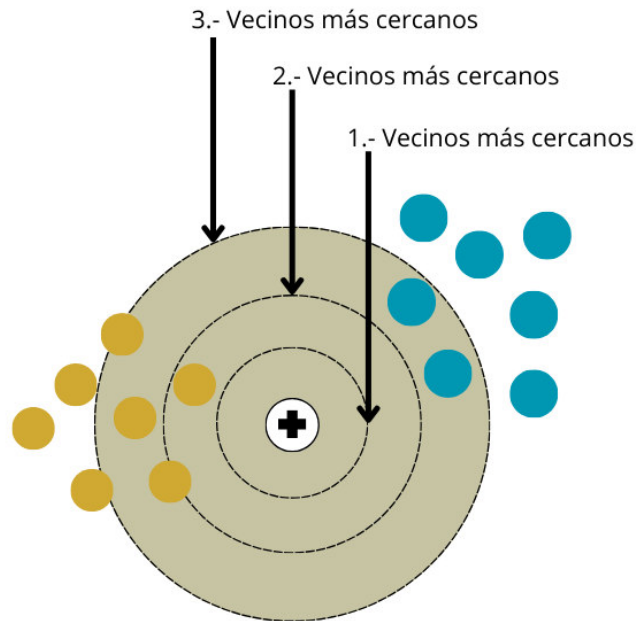


Figura 12. Ejemplo de K-vecinos más cercanos

### 3.8.5 Bosque Aleatorio

El algoritmo del bosque aleatorio (BA) o bosque de árboles de decisión se entrena en múltiples árboles de decisión impulsados por subconjuntos de datos ligeramente diferentes, se denomina simplemente como una colección de árboles y cada árbol es diferente entre sí. Construye múltiples árboles de decisión, utilizando divisiones binarias en variables predictoras para determinar las predicciones de resultados. El bosque aleatorio es parte de los métodos de conjuntos familiares que toman el árbol de decisión como un predictor individual, son basados en los métodos de embolsado, aleatorización de salidas y subespacio aleatorio que excusan el impulso. Los Árboles de decisión (AD) fueron propuestos formalmente en 2001 por Leo Breiman y Adèle Cutler, forman parte de las técnicas de aprendizaje automático. Este algoritmo combina los conceptos de subespacios aleatorios y "embolsado", es un algoritmo de aprendizaje supervisado, se utiliza en problemas de regresión y clasificación.[73], [74], [75]. El objetivo del árbol de decisión es crear dividir los datos en pequeños subconjuntos de tal manera que cada subconjunto tenga una clase dominante, ya sea clase0 o clase1.

Los árboles de decisión se utilizan predominantemente en la comunidad empresarial para trabajar en tipos de problemas de segmentación de clientes. La mayoría de los equipos de marketing de diversos sectores verticales ejecutan sus campañas de marketing sólo después de dividir a los clientes en subconjuntos lógicos, ya que una talla única no sirve para todos. Los equipos de marketing deben ejecutar estrategias de marketing más personalizadas para cada uno de los segmentos de clientes para obtener los mejores resultados y el retorno de la inversión de una campaña. Por ejemplo, si estamos ejecutando una campaña de servicio de mensajería corta (SMS) para los teléfonos móviles de nuestros clientes para vender un producto en particular, enviar el mismo contenido de SMS a toda la población puede no ser una buena idea. Puede haber una variedad de clientes en los datos; algunos de ellos pueden ser hombres, otras mujeres, algunos estudiantes, algunos profesionales en activo y algunos de ellos pueden incluso ser personas jubiladas. Los árboles de decisión son los más utilizados para trabajar con este tipo de problemas empresariales de segmentación de clientes. [39]

Debido a que el bosque aleatorio es la agrupación de varios árboles de decisión, es decir, en un bosque aleatorio se podrían encontrar 10 árboles de decisión esto con el fin de minimizar los errores. En este caso se pondrá un ejemplo del funcionamiento de un árbol de decisión, para entender la idea que se aplica para varios árboles. En el contexto de la toma de decisiones gerenciales, se presenta un

escenario hipotético en el que un gerente de una fábrica de camisetas se enfrenta a la decisión de adquirir una o dos máquinas de estampado. Cada máquina tiene un costo de \$100,000 y tiene la capacidad de producir 1,000 camisetas al día. La decisión se complica por la incertidumbre en torno a la demanda futura. Si el gerente opta por comprar solo una máquina y la demanda supera su capacidad de producción, se enfrentará a la pérdida de ventas. Sin embargo, si decide adquirir dos máquinas y la demanda es menor que la capacidad de producción combinada se enfrentará a la ineficiencia de tener una máquina ociosa.

Para ayudar en la toma de decisiones, el gerente puede emplear un árbol de decisiones. El nodo de decisión inicial en este árbol sería la elección entre comprar una o dos máquinas. A partir de este nodo, se ramificarían dos posibles caminos. Cada camino, a su vez, tendría nodos de probabilidad adicionales que representan los posibles escenarios de demanda: alta o baja, como se puede observar Figura 13.

El árbol de decisiones permite al gerente visualizar las posibles consecuencias de cada decisión y, por lo tanto, tomar una decisión más informada. Aunque este es un ejemplo simplificado, ilustra el valor de los árboles de decisiones como herramienta para la toma de decisiones bajo incertidumbre.

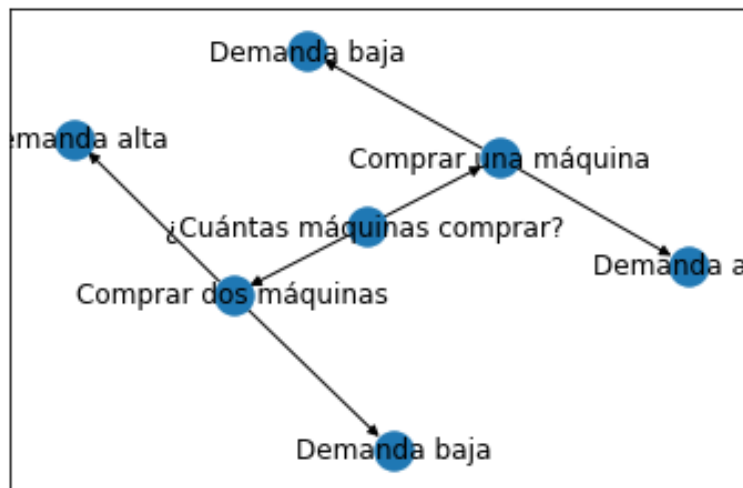


Figura 13. Ejemplo de un árbol de decisión

## 4 Metodología

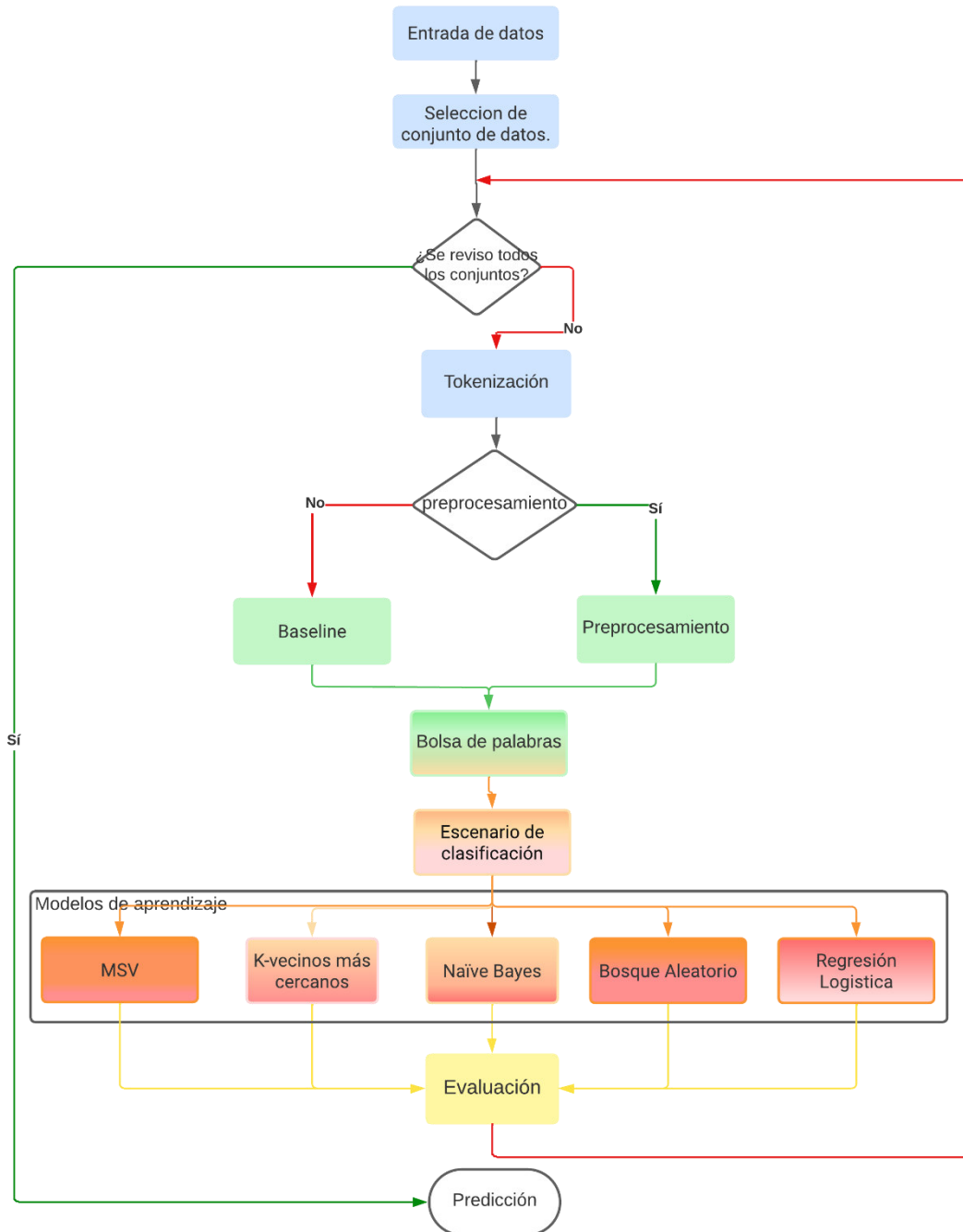


Figura 14. Diseño de algoritmo de procesamiento.

En la Figura 14 se muestra el esquema que se llevó a cabo para realizar el proceso de creación del sistema clasificador, con el fin de lograr tener una exactitud superior

o igual a lo reportado en el estado del arte. Se inicia con la entrada de datos, que son un conjunto de valores de las variables del dataset, las cuales se seleccionan como entradas. Posteriormente, generamos conjuntos de datos a trabajar que entraran en un ciclo para evaluar por cantidades los cambios que generan, si el ciclo apenas inicia, el siguiente paso será, generar nuestros tokens para poder trabajar con los datos. Después, se consideran 2 posibles caminos, en el primero se implementa una etapa de preprocesamiento y en el segundo camino se consideran los datos en crudo sin realizar preprocesamiento, a los cuales denominamos “baseline” o punto de partida. A la salida de cualquiera de los dos caminos, se procede a la creación de una bolsa de palabra la cual será analizada en los escenarios de clasificación, los cuales tiene como objetivo determinar cómo se llevaran a cabo los experimentos y que cantidades se usaran para entrenar a los modelos de aprendizaje automático (k- vecinos más cercanos , regresión logística, Naïve-Bayes, máquinas de soporte vectorial y bosque aleatorio), estos modelos con la configuración establecidas en el escenario de clasificación pasaran por medio de sus fórmulas estadísticas para poder aprender de los conjuntos de datos para poder predecir datos, estas nuevas predicciones serán las que serán evaluadas en la desempeño del sistema, esto nos sirve para evaluar una perspectiva sobre el rendimiento de nuestro modelo y nos ofrezcan una visión integral de su eficacia, finalmente una vez evaluado pasaremos a mostrar nuevos datos al sistema para predecir su categoría. En las siguientes subsecciones se describirán con mayor profundidad cada uno de los pasos de la metodología para entender mejor el proceso y los datos que se usaron para este trabajo, iniciando con la entrada de datos.

## 4.1 Entrada de datos

En esta sección, se selecciona el “dataset” con el que se trabajará, para este caso es “large movie review dataset” el cual contiene 50 mil opiniones, dividido en 2 secciones positivos y negativos, este tipo de datos se utiliza para el aprendizaje supervisado, debido a que se encuentra etiquetado por expertos en el área [76]. Un ejemplo de una entrada de datos se puede observar en la Figura 15 donde entra una opinion positiva y negativa, pasa por un proceso para que al final nos prediga que categoría pertenecen.

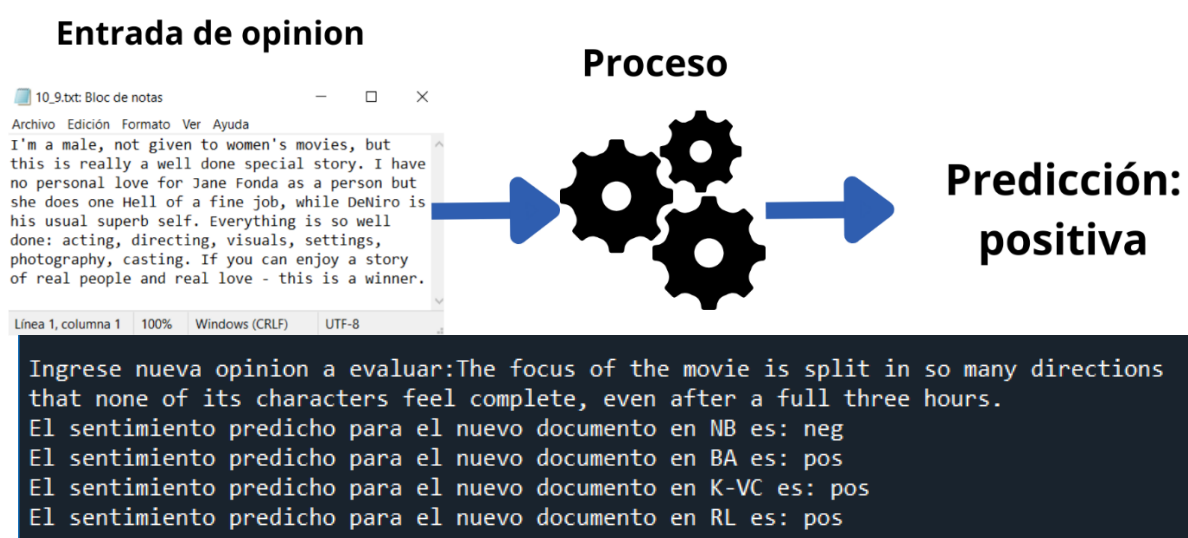


Figura 15. Ejemplo de entrada de datos con salida esperada.

## 4.2 Selección de datos

En esta sección nos sirve para generar conjuntos de datos de opiniones balanceadas por ciclo para poder observar y medir la cantidad de opiniones con las que se puede percibir cambios que afecten el desempeño de los algoritmos solo con el hecho de cambiar la cantidad de opiniones.

## 4.3 Tokenizar

La tokenización de texto consiste en segmentar el texto en oraciones y palabras especificando las unidades lingüísticas básicas: palabras, números y puntuaciones, estas son normalmente separadas por espacios en blanco, al producto se le denomina token [2], [77]. Un ejemplo de esto se puede apreciar Figura 16, en el cual tenemos una opinión hipotética acerca de la película de Avatar, esta opinión es la de algún usuario que se encuentra comentando en IMDb, el proceso de tokenizar provoca que este párrafo sea separado en palabras, y estas palabras se encuentran ahora en un espacio guardado, esto lo podemos conocer como vector, el término “vector” se utiliza porque esta representación tiene muchas propiedades matemáticas útiles que provienen del álgebra lineal.

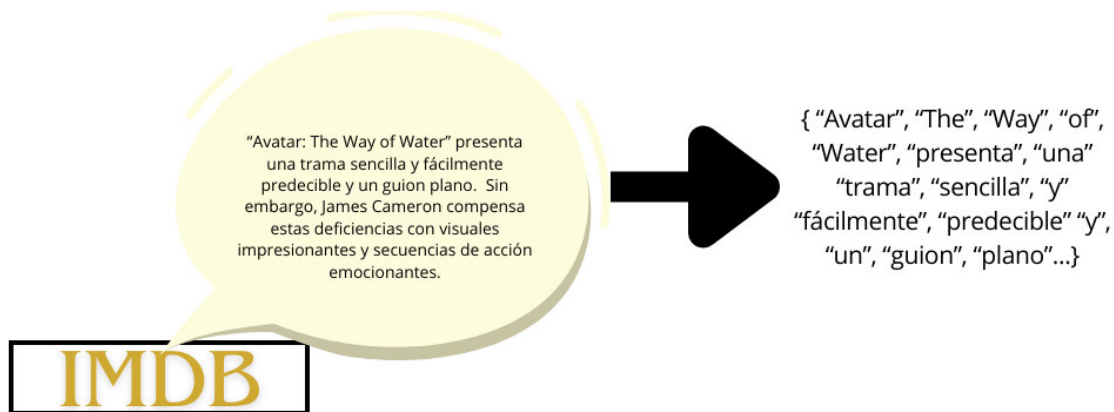


Figura 16. Ejemplo de tokenización

Es por ello por lo que, se empieza a trabajar con vectores de características, porque ahora tendrán un lugar específico dentro de nuestra representación, como se muestra en la Figura 17 con la cual podremos trabajar de manera matemática y así poder identificar los elementos de la opinión.

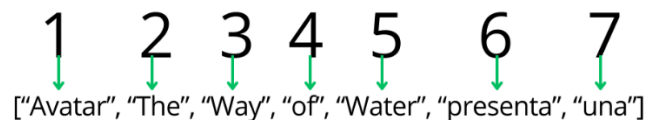


Figura 17. Ejemplo de vector de carácter

## 4.4 Preprocesamiento

El preprocesamiento puede implementarse para lograr diferentes fines y acondicionar de cierta manera la información, por ejemplo: se pueden eliminar las palabras de paro, números y acentos; también se puede realizar una transformación de las palabras a minúsculas, para reducir la dimensionalidad del vector de características, este proceso se realizará en los diferentes escenarios que presentaremos.

Una de las etapas iniciales en el preprocesamiento es la eliminación del “ruido”, el cual consiste en eliminar caracteres que interfieran en el análisis del texto. Algunos de los elementos que más comúnmente se eliminan son los siguientes:

- Links HTML.



- Marcas de HTML (“tags”).
- Expresiones entre corchetes.
- Caracteres no ASCII.
- Nombres de usuario de Twitter.
- Espacios en blanco.
- Caracteres especiales como “\”, “[”, “]”, etc. [62]

#### 4.4.1 Normalización

La normalización sirve para que los tokens tengan un formato similar, con el fin de eliminar características que a su vez permitan reducir las dimensiones de las matrices con las que se trabajará. En la Figura 18, se presenta un ejemplo de cómo se vería un conjunto sin normalizar, que, aunque sean las mismas palabras para una computadora sería completamente distintas y las tomaría como diferentes. Consecuentemente, el analizar estas palabras “semejantes” tendrá un costo computacional más costoso y complejo, es por eso por lo que se aplican técnicas para normalizar los datos. Existen diferentes tipos de normalización, por ejemplo: eliminación de palabras de paro, quitar acentos, lematización, transformación de palabras a minúsculas, entre otros, para este estudio se tomó en cuenta solo 4 procesos: eliminación de palabras de paro, quitar acentos, lematización y transformación de palabras a minúsculas.

| Archivo 1  | Archivo 2  |
|--|--|
| <pre>{ "Avatar:", "The", "Way", "of",   "Water.", "presenta", "una"   "trama", "sencilla", "y"   "fácilmente", "predecible" "y",   "un", "guion", "plano."...}</pre> | <pre>{ "avatar", "the", "way", "of",   "water.", "presenta", "una"   "trama", "sencilla", "y" "fácil",   "predecir" "y", "un", "guion",   "plano."...}</pre> |

Figura 18. Ejemplo sin normalizar

#### 4.4.2 Eliminación de acentos

En este proceso se elimina la puntuación de los acentos debido a que los métodos de aprendizaje marcan como diferente token las palabras con y sin acento, por

ejemplo “información” a “información” son palabras diferentes. Esto implica que se crean vectores y matrices de dimensiones más grandes.

#### *4.4.3 Lematización*

Algunas palabras pueden estar expresadas en diferentes tiempos de conjugación de los verbos (pasado, futuro, presente, etc.), o formas como si está en plural o singular pasan a una su forma verbal base, por lo que para simplificar su análisis se usa el proceso de lematización con el cual se convierten estas palabras en su forma base. Para esto se utiliza un vocabulario y un análisis morfológico (estructura de la palabra) de las palabras para devolver la forma básica o de diccionario de la palabra [64].

#### *4.4.4 Eliminación de mayúsculas*

En este proceso se cambian todas las letras a minúsculas (lowercase), porque para los métodos de clasificación no es lo mismo la palabra “comer” que “COMER” o que “Comer”, estas palabras son tratadas como si fueran totalmente distintas, sin ningún tipo de relación entre ellas. Por lo que puede resultar útil homogenizar las palabras y usar únicamente minúsculas para vincularlas únicamente con un significado, sin tener en cuenta la forma de cada uno de sus caracteres [78].

#### *4.4.5 Eliminación de palabras de paro*

Existe un conjunto de palabras que, aunque son necesarias para construir oraciones con sentido, carecen de información que ayude a determinar la polaridad de los textos en los que se encuentran, estas son llamadas las palabras de paro o “stop words” [78] son palabras del lenguaje que no poseen riqueza semántica, y solo agregan complejidad al análisis debido a que aumentan el tamaño de las matrices [64]. Las palabras con las que se trabajó se encuentran en el Apéndice A, esto ayuda mucho en los procesos porque al disminuir un vector de características de 100 a 50, provoca más rapidez de procesos y mayor fluidez al cargar la información, debido a que al recolectar los vectores se forma una matriz que puede quedar como Figura 19.

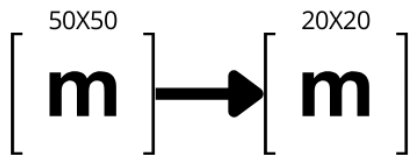


Figura 19. Reducción de matriz

En la Figura 20 se muestra el resultado después de realizar las etapas de preprocesamiento, en el cual tomando como ejemplo la opinión de avatar, se pudo reducir de 16 a 12, aunque esto a primera vista se considere poco, es una gran diferencia para el sistema debido a que por cada experimento que se realice con estos pasos realizara un trabajo con una dimensión de  $12 \times N$  siendo N la cantidad de documentos, si solo fuera 1 sería de  $12 \times 1$  pero si son 200 documentos se tendría una matriz de  $12 \times 200$ , esto implica una gran reducción de tamaño porque en algebra lineal no es lo mismo trabajar con una matriz de dimensión de  $16 \times 200$  que con una de  $12 \times 200$ .

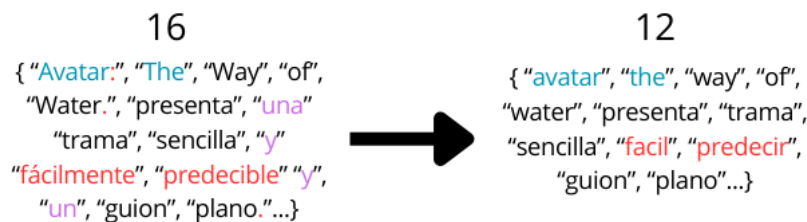


Figura 20. Ejemplo de normalización

## 4.5 Bolsa de palabras

La creación de una bolsa de palabras es una técnica de vectorización que pretende generar una representación de valores numéricos que indican el número de ocurrencias de palabras (o grupos de palabras) en un texto. Se denomina "bag" (bolsa) porque no hay ninguna ordenación u organización de las palabras utilizadas [62].

Se inicia al efectuar la evaluación de la relevancia de un término en un documento a través de la métrica conocida como frecuencia del término (tf, term-frecuency). Para llevar a cabo este proceso, se requiere la creación de una matriz documento-término (dtm), en la cual las filas representan las ocurrencias del término en un texto

determinado, mientras que las columnas representan las palabras presentes en dicho texto. Esta matriz se emplea con el propósito de identificar y extraer las palabras que poseen una relevancia superior en el contexto analizado [79].

## 4.6 Escenario de clasificación

En el escenario de clasificación, se implementó un método de muestreo aleatorio basado en dos parámetros. El primer parámetro es el porcentaje del total de datos que se utilizarán como datos de prueba. El segundo parámetro es el porcentaje de datos que se utilizarán para entrenar al algoritmo, el cual se tiene como fin tener equiprobabilidad, es decir que, utilizan métodos que buscan que todos los sujetos de una población tengan la misma probabilidad de aparecer y formar parte de la muestra [80]. En este sentido, para este proyecto usaremos un 80% de los datos para entrenamiento y el restante 20% se utilizará para el conjunto de prueba.

## 4.7 Modelos del aprendizaje automático.

En el proceso de aprendizaje automático, la selección del modelo de entrenamiento es un paso crucial. Los métodos que se han seleccionado para este estudio son Naïve Bayes, K-vecinos más cercanos, Bosque Aleatorio, Regresión logística y Máquinas de soporte vectorial. Cada método tiene sus fortalezas y debilidades, y la elección dependerá de la naturaleza de los datos y del problema que se esté abordando, en este caso, la clasificación de texto.

## 4.8 Evaluación del Desempeño del Sistema

Una herramienta útil para analizar el desempeño de un sistema es la matriz de confusión, la cual permite visualizar mediante una tabla de contingencia la distribución de errores cometidos por un clasificador. La matriz de confusión para el caso de dos clases tiene la siguiente apariencia, como se puede apreciar en la Figura 21 [81].

## MATRIZ DE CONFUSIÓN

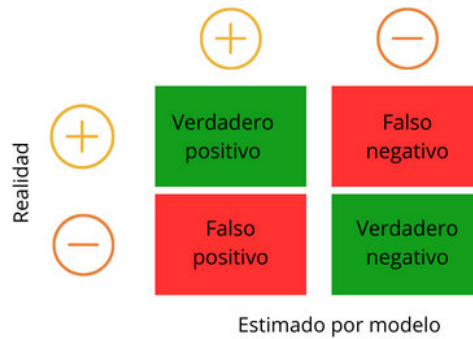


Figura 21. Ejemplo de matriz de confusión

La evaluación del desempeño de la clasificación de los sistemas implementados se realizará utilizando métricas obtenidas de la matriz de confusión como lo son la precisión, la exactitud, el recall (también conocido como sensibilidad) y el f1-score. Cada una de estas métricas proporciona la evaluación *desde* una perspectiva única sobre el rendimiento de nuestro modelo y, en conjunto, pueden ofrecer una visión integral de su eficacia.

A continuación, se describen las métricas que se emplearán en el presente trabajo de tesis. La exactitud ( $E$ ) se define como la proporción de predicciones que el modelo realiza correctamente comparada con el número total de elementos presentados. Se puede expresar con la Ecuación (8). [64], [77]:

$$E = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (8)$$

donde  $TP$ ,  $TN$ ,  $FP$  y  $FN$  representan los verdaderos positivos (true positive), los verdaderos negativos (true negative), los falsos positivos (false positive) y los falsos negativos (false negative), respectivamente.

El recall ( $R$ ) es el número de observaciones en el conjunto de prueba que el clasificador ha etiquetado correctamente ( $TP$ ) del total de registros en el conjunto de prueba que están etiquetadas para una clase particular ( $TP + FN$ ) y se puede expresar con la Ecuación (9) [82]:

$$R = \frac{TP}{TP+FN} \quad (9)$$

La precisión ( $P$ ) se define como la relación entre las muestras positivas clasificadas correctamente ( $TP$ ) y el número total de muestras pronosticadas como positivas ( $TP + FP$ ), esto se puede representar por medio de la ecuación (10) [83]:

$$P = \frac{TP}{TP+FP} \quad (10)$$

El  $F1 - score$  se define como el promedio armónico entre el recall y la precisión. Es utilizado cuando no se puede elegir entre precisión y recall, por lo que gestiona un equilibrio entre ambas métricas, esta métrica se puede expresar con la ecuación (11) [83], [84]:

$$F1 - score = \frac{2PR}{P+R} \quad (11)$$

## 4.9 Predicción

Este punto sirve para evaluar las opiniones completamente nuevas, es decir, datos que el sistema no ha evaluado y poder determinar a qué categoría pertenecen (positiva o negativa).

---

## 5 Implementación del Clasificador

---

### 5.1 Conjunto de datos

Partiendo de las 50,000 opiniones en inglés que contiene el dataset, utilizamos la carpeta de datos de entrenamiento (train) para poder entrenar al sistema. Esta carpeta de entrenamiento posee 25,000 opiniones, ordenadas del 0 hasta 12499 en cada categoría (positiva y negativa). (<https://ai.stanford.edu/~amaas/data/sentiment/>) [76].

Un ejemplo de las opiniones que contiene el dataset se muestra en la Figura 22, el cual en la parte superior tiene un identificador, en este caso se puede observar un “1”, es decir, que es la opinión uno y seguido de un valor separado por un guion bajo que nos indica la puntuación, en el ejemplo, se puede observar un “1” este valor es dentro de la escala de 1 a 10, por lo que 1 es lo más bajo posible sacar, esto nos indica qué, esta opinión fue etiquetada como negativa y se puede leer que es una crítica de la película “Stanley e Iris” de 1990, esta crítica fue extraída del sitio web IMBd, y por ende podemos encontrarla en ese sitio.

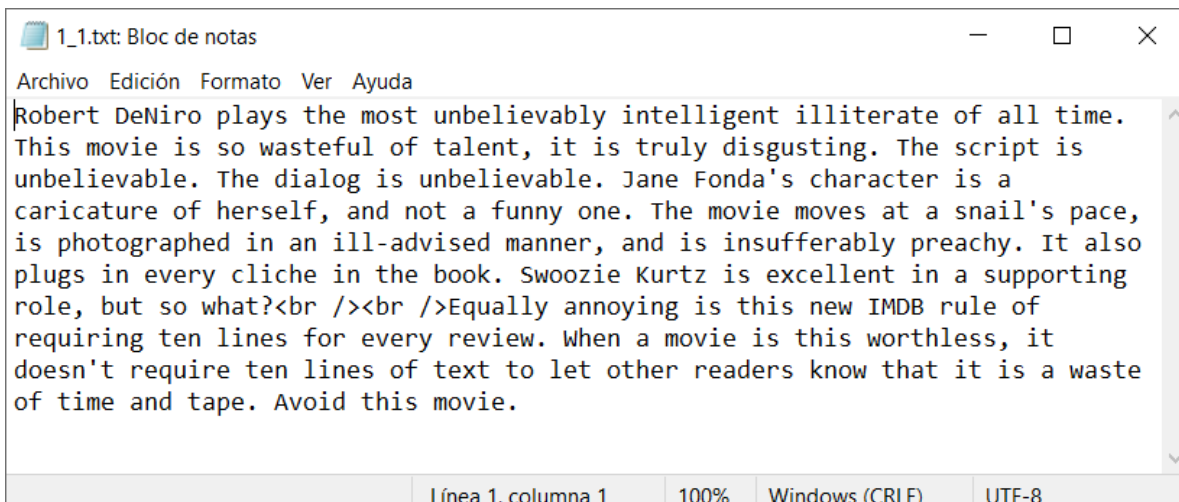


Figura 22. Ejemplo de opinión negativa

En otro ejemplo podemos observar la opinión de otro usuario en la Figura 23, etiquetada como positiva, este proceso se puede saber rápidamente porque cómo se explicó con anterioridad, al observar el número que se encuentra seguido del guion bajo se muestra el “9” significado que tiene nueve de puntuación, en esta opinión hablan de la serie de TV “Bromwell High” del 2005.

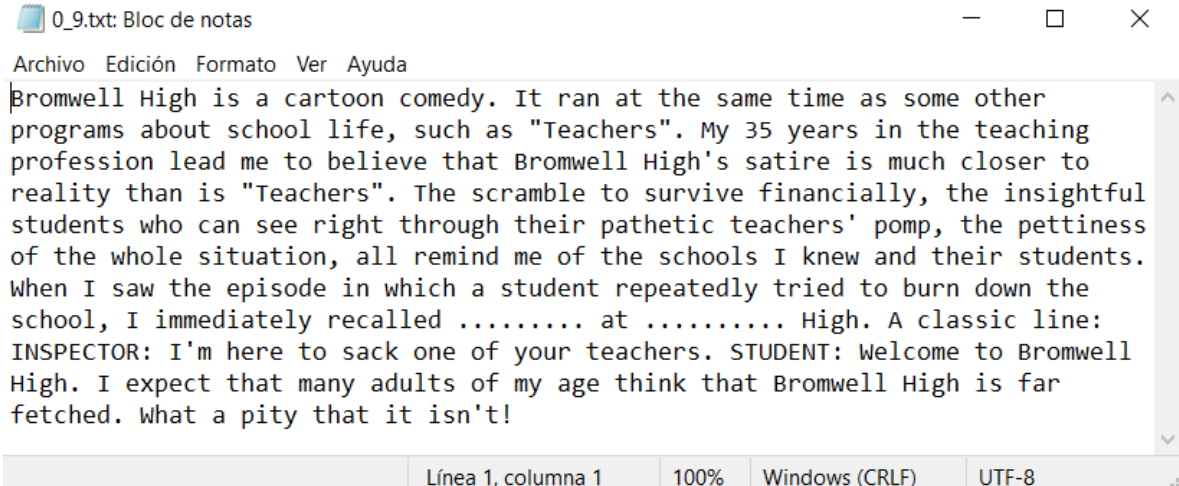


Figura 23. Ejemplo de opinión positiva

Estos ejemplos nos sirven para entender el contenido de cada documento guardado en el dataset, y saber qué tipo de información se trabaja, al hacer un análisis de opinión, nos centramos en la expresión en general por entidad (por película) y se analizan las palabras que se usan para luego poder predecir nuevas.

## 5.2 Configuración

Las configuraciones que se realizaron en un inicio para analizar qué métodos estaban generando mejores resultados son los siguientes pasos:

Primer paso en el proceso de evaluación se determinó que el número de ciclos de información que se llevaran a cabo sería de 10 ciclos. Posteriormente, es imperativo establecer un punto de partida (baseline), en este es esencial mantener una condición específica para poder realizar una medición efectiva de los 5 modelos de aprendizaje automático propuestos, en donde los datos no han sufrido ninguna alteración. De esta manera, se puede establecer si existe una mejora en las métricas de precisión, exactitud, recall y f1-score, este procedimiento permite evaluar de forma objetiva y precisa los modelos en cuestión.

El proceso de evaluación de 10 ciclos nos sirve para medir el impacto que genera introducir diferentes cantidades de opiniones, para así poder determinar con qué cantidad es mejor trabajar, para cada etapa y modelo.

Después es importante mantener la cantidad de ciclos para poder hacer una comparativa donde el cambio solo evalúe el tipo de pre-procesamiento para cada



modelo, esto con el fin de obtener la configuración idónea de cantidad de opiniones adecuada con o sin pre-procesamiento.

### 5.2.1 Punto de partida (baseline)

Tendremos la evaluación de nuestro baseline contra los 4 tipos de preprocesamiento por separado. Después realizamos la combinación de preprocesamiento para evaluar cuál configuración es mejor, las combinaciones de preprocesamiento realizadas son: eliminación de acentos y eliminación de mayúsculas, eliminación de acentos y eliminación de palabras de paro, eliminación de mayúsculas y eliminación de palabras de paro, por último, evaluamos la combinación de 3 preprocesamientos en la que consistía la siguiente combinación: eliminación de acentos, eliminación de palabras de paro y eliminación de mayúsculas.

## 5.3 Proceso

En esta sección se explicará por pasos el proceso que realiza el sistema clasificador, con un ejemplo sencillo, debido a que la cantidad de datos que se maneja son muy grandes y no se logra apreciar el proceso de manera sencilla.

### 5.3.1 Proceso de la entrada de datos a creación de bolsa de palabras

Suponiendo que tenemos un conjunto de datos de 3 documentos, que contienen la siguiente información en la Figura 24.

```
# Lista de documentos de texto
documentos = ['Este es el primer documento.',
              'Este documento es el segundo documento.',
              'Y este es el tercer documento. ¿Es este el primer documento?']
```

Figura 24. Ejemplo de documento

El primer paso es tokenizar, este proceso lo realizamos con la función llamada `CountVectorizer()` de la librería `sklearn` como se muestra en la Figura 25. Función

de tokenización. En esta etapa, lo único que logramos es separar en palabras las oraciones. Con esta función se le puede indicar qué tipo de preprocesamiento se desea realizar, para lo cual solo se agregan los argumentos correspondientes. Para la prueba de baseline no se requiere preprocesamiento por lo que no se pasa ningún argumento en la función.

```
# Crear el CountVectorizer
vectorizer = CountVectorizer()
```

Figura 25. Función de tokenización

Ahora procedemos a realizar la transformación de nuestras palabras en datos que la computadora puede entender para trabajar con ellos, por lo que transformamos los datos en números por el método de matriz documento-término, para crear una matriz en donde se encuentren las palabras por documento. Para ejecutar este proceso se emplea la función `fit_transform()` (Figura 26).

```
# Ajustar y transformar los documentos
X = vectorizer.fit_transform(documentos)
```

Figura 26. Código para crear matriz documento término.

Para que quede con mayor claridad el proceso de documento-término lo que se realiza es representar en las filas las ocurrencias del término en un texto determinado, mientras que las columnas representan las palabras presentes en dicho texto, es decir si tenemos agrupadas todas las palabras únicas de los 3 documentos tendríamos un vector así: ['documento' 'el' 'es' 'este' 'primer' 'segundo' 'tercer'], en este orden se contará cuántas veces aparece por cada documento, entonces si tenemos en el primer documento: 'Este es el primer documento.', la palabra "documento" tiene una palabra igual en el texto por lo tanto se suma 1, si consideramos que los documentos son las filas de una matriz: tendríamos que en el espacio (0,0) sería igual 1, o viéndolo desde la perspectiva de que solo es un vector sería en el espacio (0) igual a 1, este proceso sigue por cada palabra en cada documento generando una matriz de  $7 \times 3$ , como se puede observar en la Figura 27, estos datos son nuestra bolsa de palabras con los cuales ya podemos trabajar.

```

Bolsa de palabras:
[[1 1 1 1 1 0 0]
 [2 1 1 1 0 1 0]
 [2 2 2 2 1 0 1]]

```

Figura 27. Ejemplo de bolsa de trabajo

Una forma de representar estos datos, para tener mejor entendimiento, es como se presenta en la Figura 28, donde se puede apreciar la representación gráfica y la idea del proceso de la matriz documento-término, donde las filas son los documentos y las columnas los atributos.

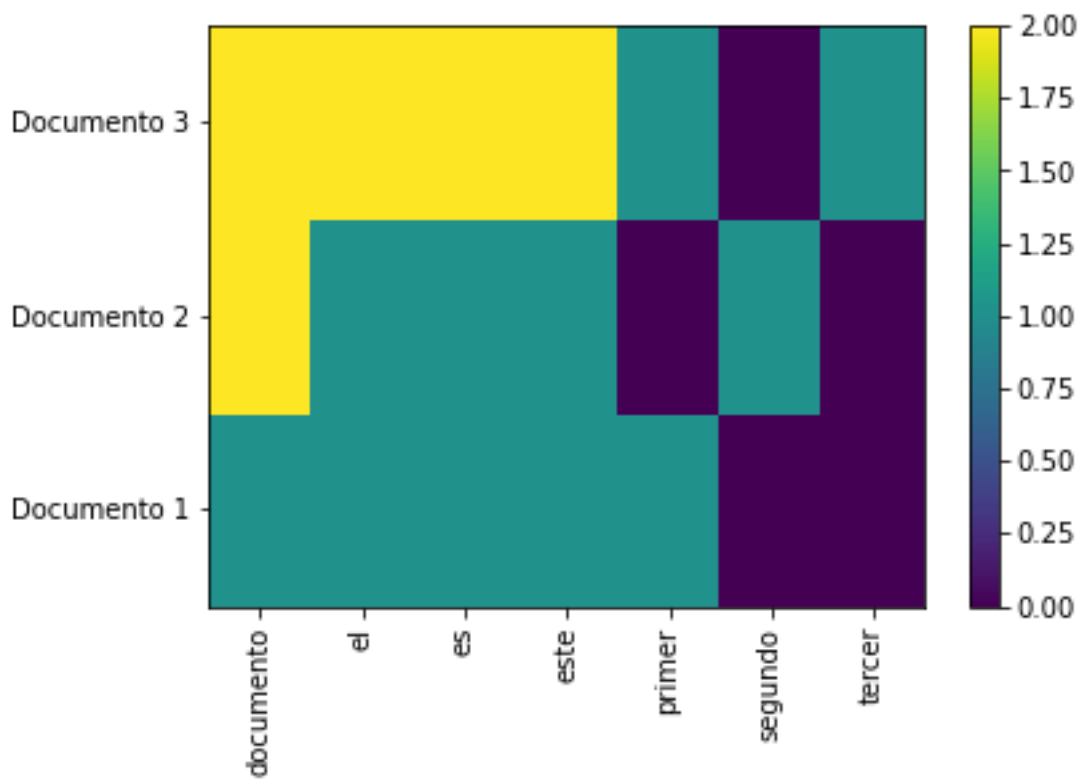


Figura 28. Ejemplo de bolsa de manera gráfica

### 5.3.2 Análisis de sentimiento

Después de tener este proceso, generamos nuestro escenario de clasificación en el cual debemos definir cuántos datos se usarán para entrenar y cuántos, para prueba, en este proceso solo trabajamos con el documento, pero como nuestro objetivo es definir si una opinión es negativa o positiva, nuestros datos deben estar etiquetados

para poder predecir nuevas opiniones, por ello ahora nuestros datos son los siguientes Figura 29, tenemos 5 opiniones y 5 etiquetas.

```
data = {
  'texto': [
    'Me encantó la película, especialmente las escenas de acción.',
    'La actuación fue increíble. Definitivamente la recomendaría.',
    'Una pérdida de tiempo. No me gustó para nada.',
    'La trama carecía de sentido y los personajes no eran creíbles.',
    '¡Una de las mejores películas que he visto!'
  ],
  'sentimiento': [
    'positivo', 'positivo', 'negativo', 'negativo', 'positivo'
  ]
}
```

Figura 29. Conjuntos de datos de análisis

Ahora nuestra bolsa de palabras queda del tamaño 33X5 Figura 30,31.

```
Bolsa de palabras:
[[1 0 0 0 1 0 1 0 1 1 0 0 0 0 1 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0]
 [0 1 0 0 0 1 0 0 0 0 1 0 0 1 2 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0]
 [0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 1 1 1 0 0 0 1 0 0 0 1 0 1 0]
 [0 0 1 1 1 0 0 1 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 1 0 0]
 [0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 1 1 1]]
```

Figura 30. Ejemplo de bolsa de palabras para análisis de sentimiento

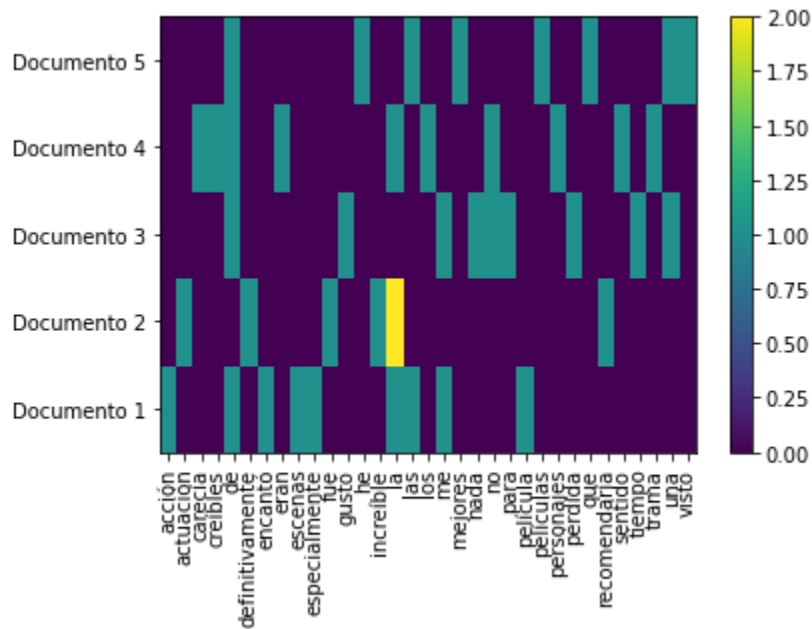


Figura 31. Ejemplo gráfico de bolsa de palabras para el análisis de sentimientos

El siguiente paso es definir el porcentaje de entrenamiento y prueba, para este caso usaremos 80% entrenamiento y 20% prueba, en la Figura 32 se observa la función que realiza este proceso, donde se indica el tamaño y los datos.

```
X_train, X_test, y_train, y_test = train_test_split(
    bolsa, # pasamos el array del texto
    data['sentimiento'], # las etiquetas que tiene
    train_size=0.80, #el tamaño del entrenamiento
    random_state=1234) #semilla de la separacion
```

Figura 32. Código para separar datos en prueba y entrenamiento.

Ahora que tenemos nuestros datos separados, pasamos esos datos a entrenar los modelos de aprendizaje automático, en este caso solo usaremos Naive Bayes como ejemplo, sin embargo, el proceso es el mismo para los cinco métodos de aprendizaje en el código solo cambian los argumentos que se indican cada método. Para el método de Naive Bayes se usa ComplementNB() como se muestra en la Figura 33, el parámetro que se pasa en este caso es Alpha, que es un parámetro de suavizado de aditivos, en este caso se pasa 0.1 por defecto, es importante este parámetro porque ayuda a manejar los casos en los que un valor característico o clase no se observó en los datos de entrenamiento, lo que podría llevar a una probabilidad calculada de 0.

```
# Inicializar y entrenar el clasificador
clfnb = ComplementNB(alpha=0.1)
clfnb.fit(X_train, y_train)
```

Figura 33. Inicialización y entrenamiento de clasificación.

Después evaluamos el clasificador con las métricas que nos interesan, en este ejemplo se muestra la evaluación de la exactitud (Figura 34).

```
# Evaluar el clasificador
train_accuracy = clfnb.score(X_train, y_train)
test_accuracy = clfnb.score(X_test, y_test)
```

Figura 34. Ejemplo de Exactitud

Para este ejemplo, el resultado fue excelente, como se puede observar en la Figura 35, debido a que el 100% de los documentos analizado fue correctamente etiquetado. En este punto es importante hacer notar que, normalmente cuando se logra un 100% de precisión el método tiende a presentar un sobre entrenamiento y que en la práctica realmente no podrá clasificar con un 100% de éxito de los casos que se le presenten. En el caso particular el ejemplo que se está presentado posee una pequeña cantidad de datos, por lo que es muy fácil determinar cada etiqueta por la falta de ejemplos.

```
Precisión en el conjunto de entrenamiento: 1.0  
Precisión en el conjunto de prueba: 1.0
```

Figura 35. Exactitud del ejemplo.

### 5.3.3 Predicción

Este proceso consiste en predecir la clase de una opinión no etiquetada, por ejemplo, en la Figura 36 se muestra una opinión nueva, que no ha sido considerada en el entrenamiento del clasificador, sin etiqueta de la categoría o clase de salida.

```
# Supongamos que tienes un nuevo documento  
nuevo_documento = ['La película fue muy decepcionante. La trama era predecible ' &  
                  ' y los personajes carecían de profundidad. Además, los efectos' &  
                  ' especiales parecían baratos y poco realistas. No la recomendaría' &  
                  ' para verla..']
```

Figura 36. Ejemplo de opinión sin etiqueta.

Para realizar una predicción los datos del nuevo documento deben de vectorizarse, y posteriormente su resultado se presenta al clasificador entrenado para que haga la predicción de a qué clase corresponde los datos del documento. En la Figura 37, se muestra el documento transformado a un vector de características, que entra el sistema a ser evaluado y el resultado se muestra en la Figura 38 donde se puede apreciar que el clasificador etiquetó correctamente la opinión, este proceso se debe hacer para cada nueva opinión que ingrese.

```
# Transformar el nuevo documento en un vector de características
nuevo_documento = vectorizer.transform(nuevo_documento)

# Usar el clasificador para predecir el sentimiento del nuevo documento
prediccion = clfnb.predict(nuevo_documento)
```

Figura 37. Código para predecir.

```
El sentimiento predicho para el nuevo documento es: negativo
```

Figura 38. Resultado del sistema.

## 5.4 Resultados del Baseline

En conclusión, en este proceso evaluamos nuestros resultados contra nuestro punto de partida para medir nuestro impacto, consideramos que este proceso sirve para identificar los aspectos más importantes de cada preprocesamiento.

### 5.4.1 Punto de partida

El primer paso para poder medir el impacto es saber cuánto se obtuvo, primero analizaremos los datos de la métrica de la exactitud que es de las más usadas en el estado del arte, debido a que nos indica “las predicciones que el modelo realiza correctamente en comparación con el número total de artículos” en la Tabla 2, podemos ver que los mejores resultados y destaca el clasificador NB obteniendo un máximo de 0.975, es decir, un 97.5% de que sus predicciones correctas ante nuevos documentos sean correctas. Este dato fue superior a lo esperado, e inclusive es ligeramente superior al 94.96% reportado en el trabajo de Manek et al. 2017. Para las demás métricas NB continúa siendo el mejor modelo con una puntuación del 97.8%, seguido de RL, MSV, y dependiendo de la métrica BA o K-VC.

Tabla 2. Punto de partida con métricas

| Método      | Máy. de E    | Máy. de P    | Máy. de R    | Máy. de F1-score |
|-------------|--------------|--------------|--------------|------------------|
| <b>BA</b>   | 0.735        | 0.837        | 0.687        | 0.739            |
| <b>K-VC</b> | 0.619        | 0.652        | 0.719        | 0.663            |
| <b>MSV</b>  | 0.77         | 0.926        | 0.803        | 0.776            |
| <b>NB</b>   | <b>0.975</b> | <b>0.978</b> | <b>0.978</b> | <b>0.978</b>     |
| <b>RL</b>   | 0.903        | 0.929        | 0.914        | 0.903            |



---

## 6 Resultados

---

Una vez que se conocen las métricas de referencia o baseline obtenidas al analizar los datos, sin ningún pre-procesamiento, con diferentes clasificadores, se procedió a repetir el proceso, pero considerando diferentes tipos de pre-procesamiento para determinar si inducen una mejora en el desempeño de los clasificadores.

### 6.1 Exactitud considerando diferentes tipos de pre-procesamiento.

A continuación, procedemos a analizar nuestros demás escenarios, comparando la misma métrica en espera de poder determinar si puede existir una mejora, en la Tabla 3 se muestra los resultados del preprocesamiento: eliminación de acentos, eliminación de palabras de paro, eliminación de mayúsculas y lematizado, y combinaciones de estos. El mejor proceso obtuvo 0.963 siendo menor al de nuestro punto de partida con 0.975, teniendo un decremento del 1.2% pero aun siendo superior al estado del arte pero al eliminar las palabras de paro baja aún más teniendo un 95% y siendo inferior por 2.5% de hecho 4 procesos de pre-procesamiento mantienen este resultado cuando se agrega este tipo, mientras que para RL el incluir el eliminación de acentos con eliminación de palabras de paro incrementa su exactitud a 92.5% siendo un 2.2% respecto a punto de partida, para MSV el incluir el proceso de eliminación de palabras de paro implica un incremento del 12% con eliminación de acentos y eliminación mayúsculas, mientras que con los demás tiene incrementos del 3.8% al 11.8%, con K-VC su incremento es del 1.7% al 3.1% y finalmente con BA el incremento va del 1.5% al 9%, siendo el ultimo el que contiene la eliminación de palabras de paro. Esto nos comprueba que el mejor resultado de todos lo encontramos en el baseline con NB. (en el apéndice B, se podrá encontrar todos los resultados). Significado de abreviaturas: Exactitud (E), Precisión (P), Recall (R), Cantidad de Opiniones (CA), Punto de Partida (PP), Lematización (Le), Eliminación de acentos (EA), Eliminación Mayúsculas (EM), Eliminación de palabras de paro (EP).

Tabla 3. Exactitud de los 4 tipos de pre-procesamiento

| Método | EA           | EA y EM      | EA y EP      | EA, EM y EP  | EM           | EM y EP      | EP           | Le           |
|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| BA     | 0.750        | 0.750        | 0.825        | 0.825        | 0.750        | 0.825        | 0.825        | 0.750        |
| K-VC   | 0.636        | 0.636        | 0.650        | 0.650        | 0.636        | 0.650        | 0.650        | 0.636        |
| MSV    | 0.808        | 0.808        | 0.890        | 0.890        | 0.808        | 0.888        | 0.888        | 0.808        |
| NB     | <b>0.963</b> | <b>0.963</b> | <b>0.950</b> | <b>0.950</b> | <b>0.963</b> | <b>0.950</b> | <b>0.950</b> | <b>0.963</b> |
| RL     | 0.892        | 0.892        | 0.925        | 0.925        | 0.892        | 0.908        | 0.908        | 0.892        |

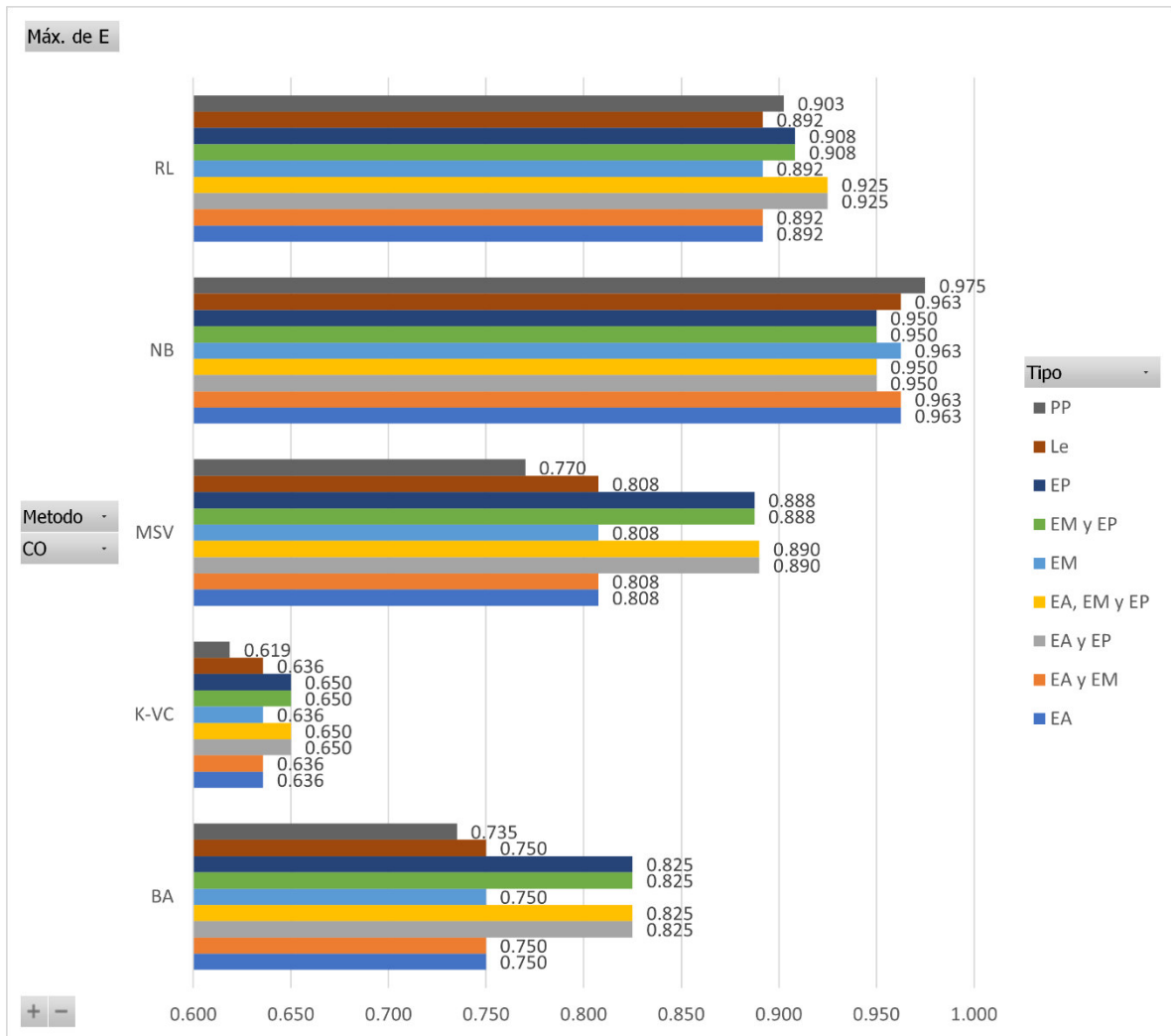


Figura 39. Gráfica de Exactitud para todos los tipos de procesos.

## 6.2 Precisión considerando diferentes tipos de pre-procesamiento

Al aplicar los preprocesamientos para la precisión vemos en la Tabla 4 que mantienen una precisión alta, en eliminación de acentos, eliminación mayúsculas, lematizado y la combinación de eliminación de acentos con eliminación de mayúsculas la mantienen en 97.7 y con eliminación de palabras de paro del 95.6 con el método de NB, siendo inferior por 0.1% y 2.2% respectivamente para el punto de partida, para BA tiene una disminución del 1.6% para eliminación de acento y la combinación con eliminación de mayúsculas mientras que para combinación de EA-AP y EA,EM con EP, sube un 2.4%, con LE y EM sube 3.2% y se ve un increíble incremento con EM y EP, y EP siendo del 13.3%, con RL su aumento solo fue cuando usamos técnicas con EP, que aumenta un 0.9%, con MSV todos disminuyen entre un 2.1% al 4.1%, finalmente con K-VC tiene un incremento del 8.1% cuando tiene EP mientras que con los demás aumenta un 0.8%.

Tabla 4. Precisión alcanzada considerando diferentes tipos pre-procesamiento

|      | EA           | EA y EM      | EA y EP      | EA, EM y EP  | EM           | EM y EP      | EP           | Le           |
|------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| BA   | 0.821        | 0.821        | 0.861        | 0.861        | 0.869        | 0.970        | 0.970        | 0.869        |
| K-VC | 0.660        | 0.660        | 0.733        | 0.733        | 0.660        | 0.733        | 0.733        | 0.660        |
| MSV  | 0.885        | 0.885        | 0.905        | 0.905        | 0.885        | 0.905        | 0.905        | 0.885        |
| NB   | <b>0.977</b> | <b>0.977</b> | <b>0.956</b> | <b>0.956</b> | <b>0.977</b> | <b>0.956</b> | <b>0.956</b> | <b>0.977</b> |
| RL   | 0.897        | 0.897        | 0.938        | 0.938        | 0.903        | 0.938        | 0.938        | 0.903        |

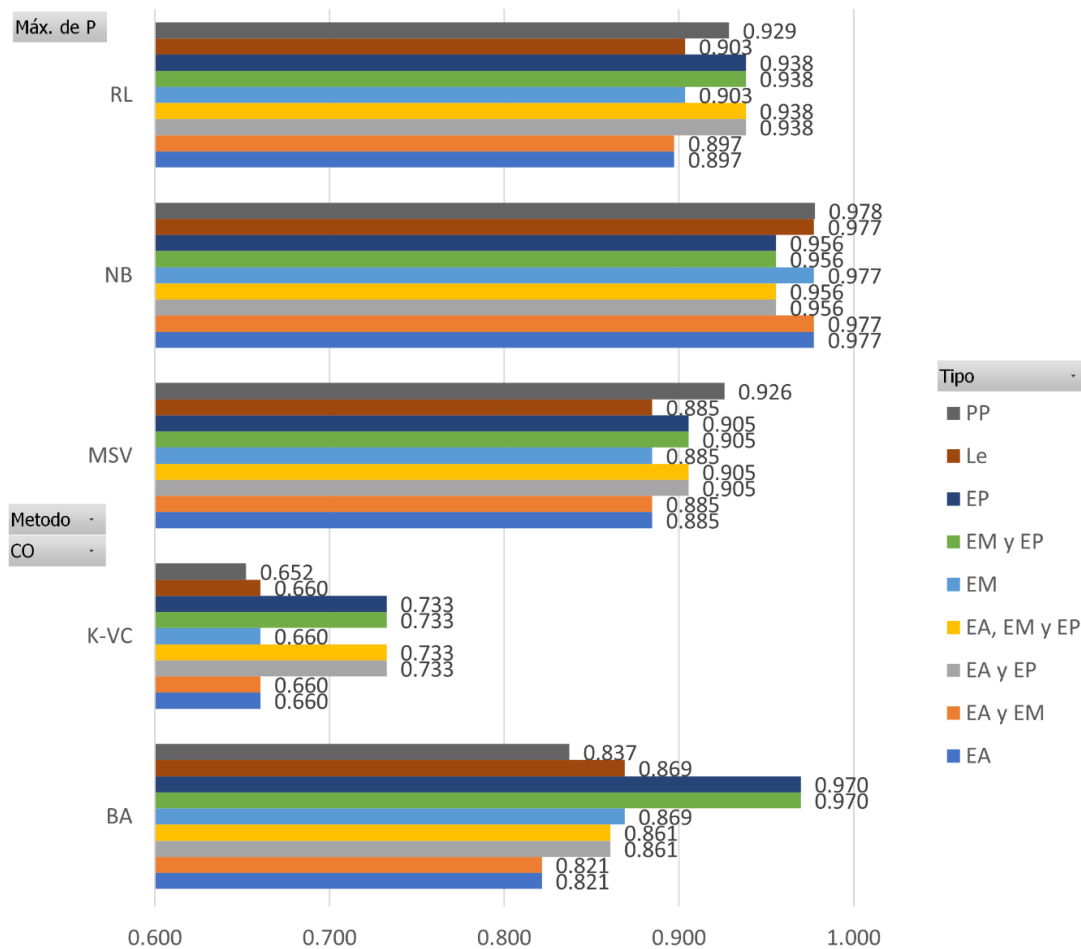


Figura 40. Gráfica de precisión para todos los tipos de procesos.

### 6.3 Recall considerando diferentes tipos de pre-procesamiento

En el contexto del análisis de sentimientos, un "recall" del 100% significa que se ha identificado correctamente todas las instancias de una determinada clase. Por ejemplo, si estamos clasificando los comentarios de los usuarios en "positivos" y "negativos", un "recall" del 100% para la clase "positiva" significaría que todos los comentarios positivos se han identificado correctamente como tales. Sin embargo, esto no garantiza que no haya habido errores en la clasificación de los comentarios negativos como positivos. En otras palabras, el "recall" se centra en la exhaustividad de la identificación correcta de una clase, pero no necesariamente en la precisión de esa identificación. En este caso MSV obtuvo el 100% para todos los tipos de pre-

procesamiento, teniendo un incremento del 19.7% respecto a su punto de partida, pero este no fue el mayor incremento en BA con combinaciones se logró tener un incremento del 25.43% cuando se hace uso de las combinaciones de EA y EP, y EA, EM y EP, lo que significa que el contener EA y EP para este modelo es muy beneficioso, en NB disminuyo con un 2.2%, con K-VC su incremento fue del 0.83% al 4.58% y finalmente el RL su incremento fue de 0.39% al 2.03%.

Tabla 5. Recall obtenido al aplicar tipos de pre-procesamiento

|            | EA           | EA y EM      | EA y EP      | EA, EM y EP  | EM           | EM y EP      | EP           | Le           |
|------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| BA         | 0.721        | 0.721        | 0.941        | 0.941        | 0.824        | 0.820        | 0.820        | 0.824        |
| K-VC       | 0.727        | 0.727        | 0.765        | 0.765        | 0.727        | 0.765        | 0.765        | 0.727        |
| <b>MSV</b> | <b>1.000</b> | <b>1.000</b> | <b>1.000</b> | <b>1.000</b> | <b>1.000</b> | <b>1.000</b> | <b>1.000</b> | <b>1.000</b> |
| NB         | 0.956        | 0.956        | 0.956        | 0.956        | 0.956        | 0.956        | 0.956        | 0.956        |
| RL         | 0.918        | 0.918        | 0.934        | 0.934        | 0.918        | 0.934        | 0.934        | 0.918        |

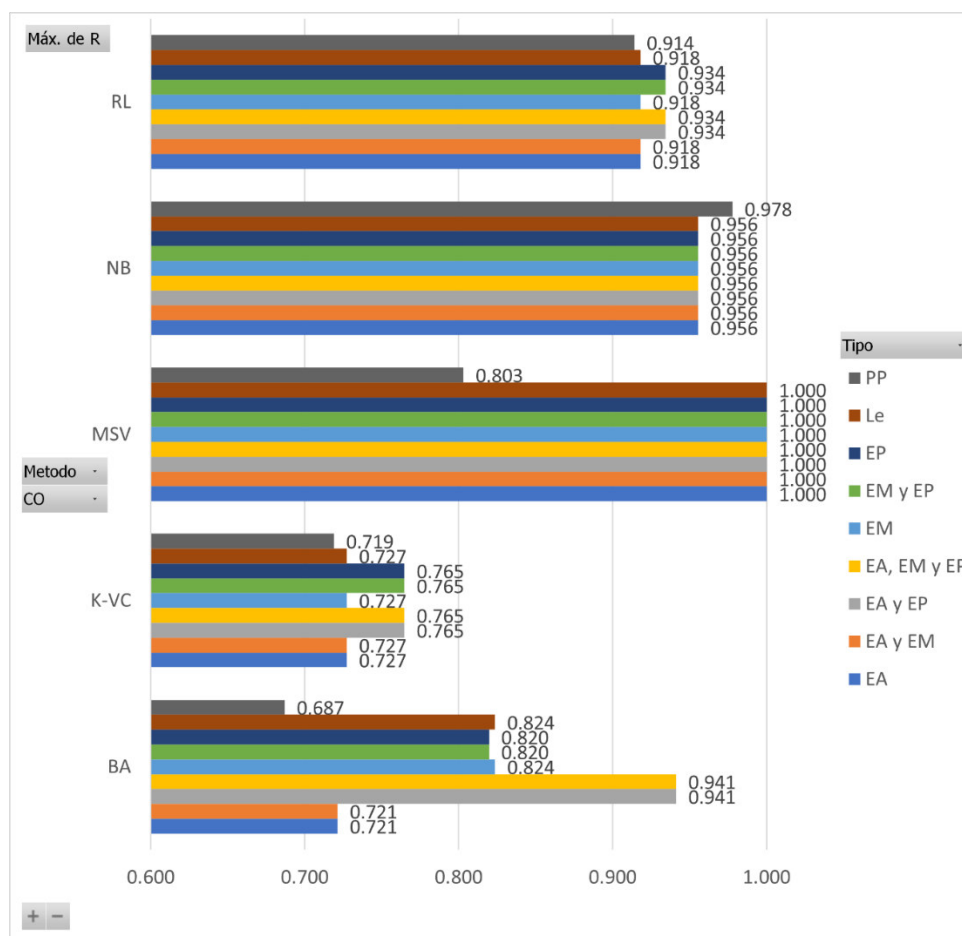


Figura 41. Gráfica de recall para todos los tipos de procesos.

## 6.4 F1-score considerando diferentes tipos de pre-procesamiento

En este caso el mejor resultado se muestra en la Tabla 6 es de NB con 96.6% que empata con los procesos de preprocesamiento: eliminación de mayúsculas, eliminación de acentos y lematizando, es decir que incluir estos procesos solo disminuye un poco en comparación al punto de partida y podemos observar que al combinar los preprocesamientos tenemos una reducción mayor respecto a lo individual y al punto de partida, el no incluir palabras de paro baja la armonía en 1.15% pero el tener un 95% es aún un resultado muy favorecedor, mientras que RL tiende a estar entre  $-0.16\%$  al  $0.93\%$ , el incremento más significativo lo vemos con MSV logrando un  $11.55\%$  cuando se realiza la eliminación de palabras de paro, también sucede este incremento para eso mismas técnicas con BA logrando un incremento del  $8.11\%$  y una pequeña mejora en K-VC cuándo se usa cualquier tipo de pre-procesamiento logrando un  $1.01\%$ .

Tabla 6. F1-score obtenidos al aplicar diferentes tipos de pre-procesamiento

|      | EA           | EA y EM      | EA y EP | EA, EM y EP | EM           | EM y EP | EP    | Le           |
|------|--------------|--------------|---------|-------------|--------------|---------|-------|--------------|
| BA   | 0.746        | 0.746        | 0.821   | 0.821       | 0.737        | 0.821   | 0.821 | 0.737        |
| K-VC | 0.673        | 0.673        | 0.668   | 0.668       | 0.673        | 0.668   | 0.668 | 0.673        |
| MSV  | 0.810        | 0.810        | 0.891   | 0.891       | 0.810        | 0.889   | 0.889 | 0.810        |
| NB   | <b>0.966</b> | <b>0.966</b> | 0.956   | 0.956       | <b>0.966</b> | 0.956   | 0.956 | <b>0.966</b> |
| RL   | 0.901        | 0.901        | 0.912   | 0.912       | 0.901        | 0.912   | 0.912 | 0.901        |

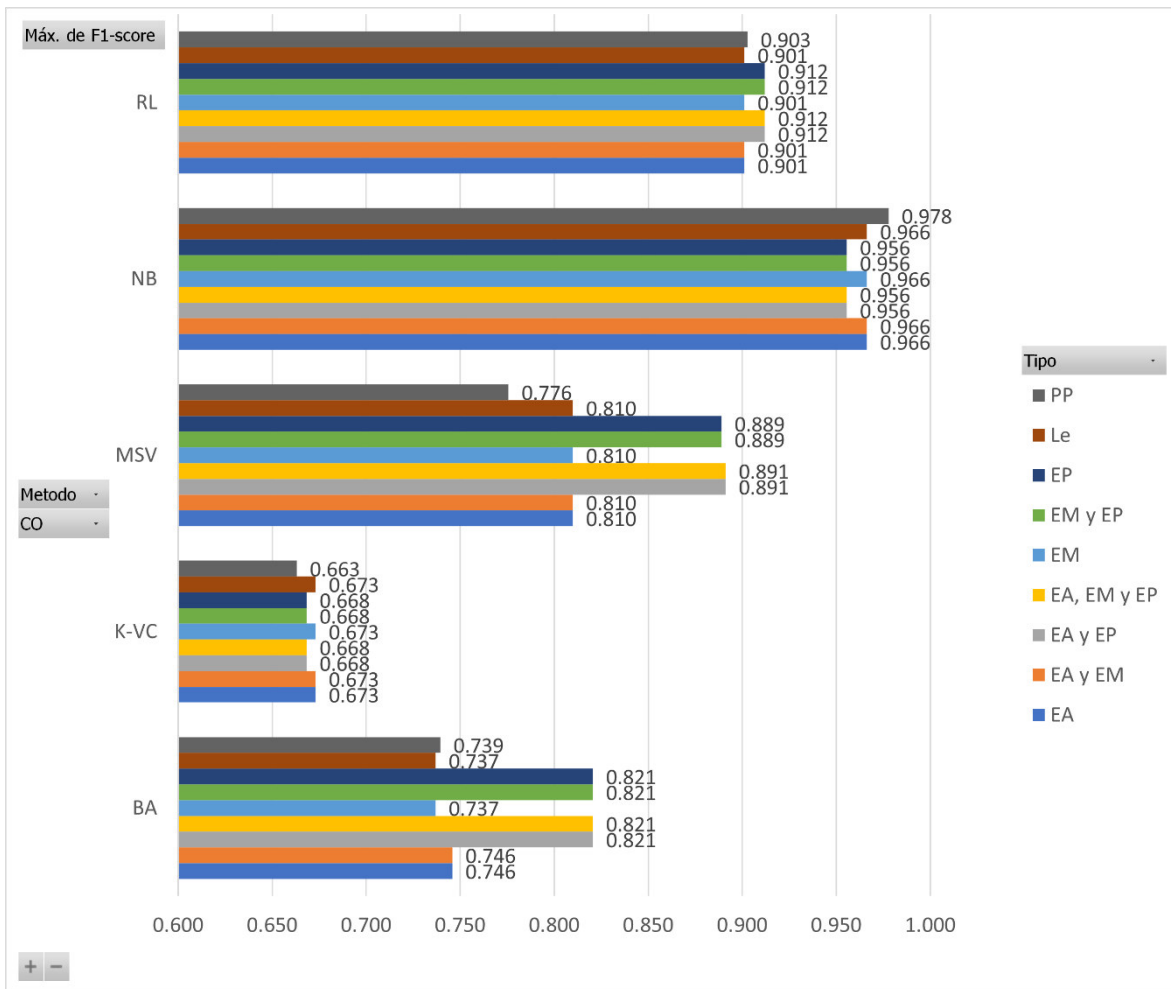


Figura 42. Gráfica de f1-score para todos los tipos de procesos.

## 6.5 Resultado final

En conclusión, los mejor resultado obtenido se observa en la Tabla 7, el principal destacado es NB en su estado base el cual utilizan solo 400 opiniones, además de, que para todas las métricas es el mejor, para interpretar los datos de la Tabla 7 nos arrojan que tenemos un modelo que predice correctamente las opiniones positivas y negativas para cada caso, y la probabilidad de que falle es muy baja, y no es necesario entrenar de más al modelo porque entre más datos se den disminuye su exactitud. Además de que este modelo garantiza una predicción del 97.5% sean correctos, otro de los efectos observados es mejor solo utilizar los datos sin pre-procesamiento, porque al utilizar etapas de pre-procesamiento puede reducir sus métricas, no en todos los casos, pero la mejor configuración es o utilizar solo nuestro baseline o usar uno de los 3 tipos de pre-procesamiento como: eliminación de acentos o eliminación de mayúsculas o lematizar, para tener un resultado

equiparable al punto de salida, mientras que para RL es mejor utilizar pre-procesamiento, específicamente eliminación de acentos con eliminación de palabras de paro, logrando un 0.925 de exactitud, mientras que para K-VC y BA, el eliminar palabras de paro genera incrementos en estos.

Tabla 7. Mejor resultado con métricas

| Método    | Exactitud    | Precisión    | Recall       | F1           | No. Opiniones | Prueba   | Tipo      |
|-----------|--------------|--------------|--------------|--------------|---------------|----------|-----------|
| <b>NB</b> | <b>0.975</b> | <b>0.978</b> | <b>0.978</b> | <b>0.978</b> | <b>400</b>    | <b>2</b> | <b>PP</b> |
| RL        | 0.925        | 0.938        | 0.882        | 0.909        | 200           | 41       | EA y EP   |
| K-VC      | 0.650        | 0.565        | 0.765        | 0.650        | 200           | 31       | EP        |
| BA        | 0.825        | 0.970        | 0.711        | 0.821        | 400           | 32       | EP        |
| MSV       | 0.890        | 0.874        | 0.909        | 0.891        | 2000          | 50       | EA y EP   |



---

## 7 Conclusiones

---

Los métodos de aprendizaje automático para la clasificación, como Naïve Bayes, K-vecinos más cercanos, bosque aleatorio, regresión logística y máquinas de soporte vectorial, son técnicas computacionales que se utilizan para predecir la categoría o clase de una entidad dada, basándose en un conjunto de características o atributos. Estos modelos se entrenan con un conjunto de datos de entrenamiento y luego se prueban con un conjunto de datos de prueba para evaluar su rendimiento. Las métricas de rendimiento, como son la precisión, la exactitud, el recall y el f1-score, se utilizan para cuantificar la eficacia del modelo en la clasificación correcta de las entidades.

En el presente trabajo se presentó una propuesta de un sistema capaz de clasificar automáticamente los sentimientos expresados en las opiniones de películas de largometraje. Dicho sistema está basado en técnicas de inteligencia artificial y ha demostrado clasificar los sentimientos con una precisión superior a la reportada por métodos similares en la literatura existente. En particular, con el modelo de aprendizaje Naive Bayes se logró obtener el mejor resultado de este sistema fue del 97.5% de exactitud, usando solo 400 opiniones, pero este modelo es considerado robusto debido a que, en precisión, recall y f1-score obtuvo; 0.978, mientras que con K-VC y RL usando solo 200 opiniones registran los mejores resultados cuando se usa EP y EA con EP, respectivamente, RL logro alcanzar una exactitud del 0.925, precisión del 0.938, recall del 0.882 y finalmente f1-score del 0.909, estos resultados son muy buenos porque 3 de 4 están arriba del 90%, significando una garantía mejor al clasificar, en K-VC se logró un incremento del 8% respecto a su punto de partida en exactitud, 6.5% en precisión, 5.9% en recall y f1-score 6.5%, con BA un incremento del 14% para la exactitud y precisión, un 15% para recall y 12% para f1-score, finalmente con MSV se incrementó un 12% en exactitud y precisión, un 11% para el recall y finalmente en f1-score 11.5%. Estos resultados destacan el potencial y la eficacia de las técnicas de inteligencia artificial en la clasificación automática de sentimientos, es de destacar que con este análisis se descubrió que tipos de pre-procesamiento le combine a cada modelo y nos mostró que para casos como NB *no es indispensable* utilizar etapas de pre-procesamiento en la implementación de un sistema de clasificación automática de opiniones de largo metraje, mientras que para los 4 restantes consideramos útil utilizar de estas técnicas de pre-procesamiento, esto demuestra que es posible diseñar metodologías que hagan uso de diferentes técnicas de inteligencia artificial, y que puedan competir contra los

reportados en la literatura, la metodología planteada supera a los reportados por la literatura demostrando que el seleccionar correctamente un método de aprendizaje nos puede garantizar una superioridad al clasificar opiniones y que el implementar etapas de preprocesamiento puede o no beneficiar en ciertas circunstancias su uso, el análisis de este proyecto nos indica que Naïve Bayes es por mucho el mejor para el análisis de opiniones y que el usarlo nos garantiza un alto rendimiento de predicción y certeza al etiquetar positivamente o negativamente las opiniones.

Finalmente, se propone trabajos a futuro sobre este sistema que podría lograr ser un sistema inteligente que por medio de la identificación de polaridad en las opiniones públicas sobre las películas se pueda crear un sistema de recomendación y ser útil no solo para hacer recomendaciones de películas, sino también para identificar tendencias cinematográficas en las redes sociales. Proporcionará un desglose detallado de las opiniones positivas y negativas, ofreciendo así una visión clara de la percepción pública para identificar tendencias en la red.

Como información adicional, este proyecto se realizó en un equipo personal, usando el lenguaje de programación de python con una GUI de Anaconda, librerías de sklearn, pandas, nltk. En la parte de hardware el equipo cuenta con un procesador i7-10750H CPU @ 2.60GHz y 2.59 GHz, 16GB de RAM, y un sistema operativo Windows 10 de 64bits.

---

## 8 Producción científica

---

Sierra Rangel, J. A., Melendez Cardiel, M. A., Arengas Acosta, J. M., & Lopez Lemus, J. A. (2022). Metodologías Ágiles en Administración de Proyectos de Software. *Academia Journals Chetumal*, 539-543.

### Metodologías Ágiles en Administración de Proyectos de Software

Ing. Jesús Andrés Sierra Rangel<sup>1</sup>, Ing. Michelle Alejandra Meléndez Cardiel<sup>2</sup>,  
Ing. José Manuel Arengas Acosta<sup>3</sup>, Dr. Jorge Armando López Lemus<sup>4</sup>

**Resumen**—Este trabajo muestra un análisis de diversas metodologías ágiles empleadas en la administración de proyectos, teniendo un enfoque en el área de proyectos de software, en el cual se busca proporcionar un marco referencial en la planificación de la gestión en el proceso del desarrollo de software, realizando una búsqueda metódica/sistemática de diferentes fuentes de investigación como lo son artículos, revistas, libros, entre otros, acerca de las metodologías ágiles para la administración de proyectos en los que se encuentran scrum, xp y kanban, las cuales se implementaron en diferentes bases de investigación científica. Se realizó un análisis y comparación de las diferentes metodologías siguiendo un criterio cualitativo para la revisión de sus respectivas ventajas y desventajas. Los resultados obtenidos muestran que la metodología scrum es la más utilizada por los desarrolladores, debido a su cantidad de recursos de gestión que facilitan la reducción de tiempos de ejecución de los proyectos.

**Palabras clave**— Metodologías ágiles, software, scrum, kanban, xp

#### Introducción

Las metodologías tradicionales de desarrollo son inflexibles y no tiene una respuesta inmediata a las solicitudes de los clientes, surgiendo las metodologías ágiles como un conjunto de prácticas que permiten la adaptación rápida de necesidades al momento del desarrollo de software. (Papadopoulos, 2015). Las metodologías ágiles surgieron en los años noventa como metodologías de desarrollo ligero, con el tiempo estas metodologías fueron cambiando el mundo de la industria del software dando a los métodos ligeros y ágiles cambios a un ritmo eficiente, evitando reducir la probabilidad de fracaso por subestimación de costos, tiempos y funcionalidad en los proyectos. (Chaouch et al., 2019; Cadavid et al., 2013).

El desarrollo de software se ha visto en un aumento debido a la productividad que se manejan en estas metodologías, mostrando flexibilidad para equipos pequeños (Mitre-Hernández Hugo et al., 2014) es por eso por lo que es necesario elegir o seleccionar las metodologías que se enfoquen en la eficiencia en proyectos con requisitos muy cambiantes, y provoquen la reducción de los tiempos de desarrollo, manteniendo la calidad del software (Patilla HJ et al., 2021).

Cada vez más empresas están adoptando métodos ágiles como una forma flexible de introducir nuevos productos de software. Una parte importante de cualquier proyecto de software es la prueba. Las pruebas ágiles pueden tener objetivos similares a las pruebas de software tradicionales, pero la estructura del equipo es diferente, los evaluadores deben respaldar la infusión de calidad en todo el equipo. (Rajasekhar, P. & Shafi, RM., 2014).

#### Metodología

La metodología que se realizó para llevar a cabo este artículo consta de 4 procesos: primero la búsqueda de información de fuentes como revistas, libros, artículos y base de datos acerca de las metodologías ágiles enfocados en scrum, xp o kanban, después de la búsqueda, la selección de artículos más actuales, para esto se recopilamos alrededor de 35 para la comprensión de las estructuras de las metodologías, las ventajas o desventajas y sus procesos, continuando con la comparación de las metodologías en base a la información recabada para la obtención del marco referencial.

<sup>1</sup> El Ing. Jesús Andrés Sierra Rangel actualmente estudiante de la Maestría en Administración de Tecnologías en la Universidad de Guanajuato, México. [jsierrarangel@ugto.mx](mailto:jsierrarangel@ugto.mx)

<sup>2</sup> La Ing. Michelle Alejandra Meléndez Cardiel actualmente estudiante de la Maestría en Administración de Tecnologías en la Universidad de Guanajuato, México [ma.melendezcardiel@ugto.mx](mailto:ma.melendezcardiel@ugto.mx)

<sup>3</sup> El Ing. José Manuel Arengas Acosta es estudiante de Maestría de Administración de Tecnologías de la Universidad de Guanajuato, estudio la Ingeniería en la Universidad Industrial de Santander, Bucaramanga, Santander, Colombia [jaramasacosta@ugto.mx](mailto:jaramasacosta@ugto.mx)

<sup>4</sup> La Dr. Jorge Armando López Lemus es docente de tiempo completo de la Universidad de Guanajuato, Yuriria, Guanajuato, México. [lopezja@ugto.mx](mailto:lopezja@ugto.mx)

Sierra Rangel, J. A., & Guzmán Cabrera, R. (2022). Clasificación automática de opiniones de películas usando . *Encuentro en Ciencia y Tecnología*, 1-7.

## Clasificación automática de opiniones de películas usando aprendizaje automático.

Jesús Andres Sierra Rangel and Rafael Guzmán Cabrera

Universidad de Guanajuato, departamento de estudios multidisciplinarios, Yuriria, Guanajuato, México  
{ja.sierrarangel, guzmanc}@ugto.mx

**Abstract.** El emitir opiniones sobre productos o servicios en foros y redes sociales se ha convertido en una tarea cotidiana. Cuando el número de opiniones es muy grande se requiere contar con herramientas automáticas que realicen esta tarea eficientemente. En este trabajo se presentan resultados obtenidos al llevar a cabo la clasificación automática de opiniones en distintos escenarios de clasificación y con diferentes conjuntos del corpus "Large movie review dataset". La idea principal de la configuración experimental propuesta fue mostrar la cantidad de opiniones que se requiere como mínimo para que el sistema de aprendizaje automático pueda aprender las características que distinguen una clase de otra. Los resultados obtenidos nos muestran que dependiendo del método de aprendizaje existe un sobre aprendizaje para métodos como SVM, KNN y Random Forest, mientras que para métodos como Naive Bayes los mejores resultados son los que deben utilizar un 70-80 de los atributos y solo en Logistic regression debe usarse en crudo para tener los mejores resultados.

**Keywords:** Large movie review dataset, svm, naïve baiyes, random forest, knn, logistic regression.

### 1 Introducción

Las tecnologías han evolucionado y se ha encontrado que la gran mayoría de información se encuentra en forma no estructura [1] se estima que entre el 80% y el 90% de los datos de las organizaciones son no estructurados, para obtener esa información es necesario procesarla y tener una estructura que se pueda analizar, para esto tenemos la minería de datos, dentro de la minería de datos, encontramos una rama llamada "text mining" o la minería de texto que [2, 3] es el descubrimiento por computadora de información nueva, previamente desconocida, mediante la extracción automática de información, en esta rama se trabaja con el análisis de sentimientos, se define [4] como minería de opinión o "opinion mining", se trata de una tarea de clasificación masiva de documentos de manera automática, que se centra, entre otras cosas, en catalogar los documentos en función de la connotación positiva o negativa del lenguaje utilizado en el mismo, [5] el enfoque común para comprender la opinión es verlo desde ambas perspectivas, es decir, positiva y negativa y el proceso se llama 'análisis de sentimiento'. [6] Las opiniones clasificadas se han identificado como tarea primordial del sentimiento.



Sierra Rangel, J. A., Guzmán Cabrera, R., & Vargas Rodríguez, E. (2023). Estudio y comparativa de algoritmos para el análisis de sentimientos en opiniones de películas. *CINERGIA UG*, 1-5.

J. Sierra et al.: Estudio y comparativa de algoritmos para el análisis de sentimientos en opiniones de películas

## Estudio y comparativa de algoritmos para el análisis de sentimientos en opiniones de películas

JESUS ANDRES SIERRA RANGEL<sup>1</sup>, RAFAEL GUZMÁN CABRERA<sup>1</sup>, EVERARDO VARGAS RODRÍGUEZ<sup>1</sup>.

<sup>1</sup>Universidad de Guanajuato, Campus Irapuato-Salamanca, Departamento de Estudios Multidisciplinarios sede Yuriria.  
ja.sierrarangel@ugto.mx  
guzmanr@ugto.mx  
evr@ugto.mx

**RESUMEN** El objetivo de este artículo es analizar los algoritmos: MSV, RL y Naïve Bayes, en base al corpus "large movie review dataset" viendo por medio de estadísticas los resultados obtenidos en pruebas realizadas por medio de un algoritmo de clasificación de opiniones de películas, descubriendo que el mejor método para realizar la clasificación de opiniones es usar NB.

**PALABRAS CLAVE** – Opiniones, sentimientos, análisis, aprendizaje, automático, películas

### I. INTRODUCCIÓN

La industria del entretenimiento actualmente es de los negocios con mayor volumen a nivel mundial, además de tener una situación social favorable que beneficia este medio desde la llegada de la pandemia que marcó un antes y después para la sociedad mundial [1], provocando un cambio en la interacción de los públicos al ser aislados, momento donde se aumentó los clientes de las plataformas de "streaming" (es la distribución digital de archivos a través de la difusión continua por una red informática [2]), creando sociedades en la red, donde cada persona publica su opinión sobre el producto al saber si le gusta o no, esto sumado a diferentes canales de información, como: redes sociales, canales dedicados, blogs, sitios web de reseñas de películas y series (Ej. Rotten Tomatoes, Internet Movie Database, Filmaffinity, etc.) provocó las creaciones de enormes cantidades de opiniones sobre un producto o servicio, la explotación de estos datos requiere de herramientas del área de una inteligencia que pueda extraer esta información [3]-[5].

Basado en el informe de "Global Digital Trends" el número de usuarios en el planeta para abril del 2023 es de aproximadamente 5,180 millones de usuarios, siendo un total del 64.6 % de la población mundial, donde en promedio pasan un total de 6 horas con 35 minutos por usuario en internet, y de esto el porcentaje de usuarios en internet de los 16 a los 64 años que desean ver algo por medio de "streaming" es del 90.1% [6].

A lo largo de los años se han establecido el alto grado de causalidad que desempeñan las reseñas y los comentarios en línea sobre el resultado de una película. Frecuentemente, las opiniones tienen una gran influencia debido a que las personas deciden que producto consumir tomando en consideración las evaluaciones y opiniones de una tercera persona [7]. Para el caso de productos como películas, se

genera una gran cantidad de datos de forma de "electronic word of mouth" (eWOM) que se refiere a la forma de definir un intercambio de información no comercial escrita. Una de las tareas que en años recientes se ha vuelto popular es el análisis de sentimientos [8], el cual involucra varias áreas como son: el procesamiento del lenguaje natural, el uso de minería de texto y el aprendizaje automático. La actividad de analizar el texto de opiniones o reseñas se trabaja con el objetivo de identificar la emoción detrás de la opinión (a favor, en contra, positivo, negativo, etc.). En este sentido, algunos autores definen como minería de opiniones, al proceso de identificar la emoción a partir de miles de comentarios de manera automática [9].

El análisis de opiniones es un tema que cada día va tomando más peso debido al creciente uso de las redes sociales, las sociedades en la red y los públicos, que producen y reproducen sus capacidades de creación de contenido, participan de manera activa en las discusiones de temas: sociales, políticos, deportivos, entretenimiento y servicios. [3]

Es por ello por lo que la actual producción científica va enfocada en la extracción de las emociones encontradas en las opiniones de la red, centrándose en un análisis de sentimiento como metodología basada en la minería de textos sobre el "big data" [1], [10].

El disponer de una fuente tan rica de información representa oportunidades para los interesados, pero también es una complicación del cual implica un análisis y una confiabilidad que brinden las herramientas dedicadas a solucionar de manera precisa la extracción de las opiniones [5].

Definir con precisión si un comentario es positivo o negativo requiere de un modelo supervisado, el cual tiene 2 características importantes: los algoritmos de clasificación a

Congreso Internacional de Energía UG, 27 - 29 de septiembre de 2023

---

## Referencias

---

- [1] J. C. Alcalde, D. Isern, A. Carles, and V. Royo, "Análisis de sentimiento de textos basado en opiniones de películas usando algoritmos de aprendizaje computacional," Universitat Oberta de Catalunya, 2021. Accessed: Jun. 29, 2023. [Online]. Available: <https://openaccess.uoc.edu/handle/10609/132328>
- [2] C. A. Arango Pastrana and C. F. Osorio Andrade, "Aislamiento social obligatorio: un análisis de sentimientos mediante machine learning," *Suma de Negocios*, vol. 12, no. 26, pp. 1–13, Dec. 2021, doi: 10.14349/SUMNEG/2021.V12.N26.A1.
- [3] Á. Pilao Rodríguez, "La industria musical, la cibernética y el 'streaming'. Un análisis de la evolución musical desde los años 80 a la actualidad," Universidad de Sevilla, Sevilla, 2020. Accessed: Jun. 29, 2023. [Online]. Available: <https://idus.us.es/handle/11441/101474>
- [4] K. Rodríguez-Díaz and Y. Haber-Guerra, "Análisis de sentimientos en Twitter aplicado al #impeachment de Donald Trump," *Revista Mediterránea de Comunicación*, vol. 11, no. 2, p. 199, Jul. 2020, doi: 10.14198/MEDCOM2020.11.2.23).
- [5] S. Belén and A. Viteri, "Análisis de sentimientos para Twitter con Vader y TextBlob," *Revista Odigos*, vol. 2, no. 3, pp. 9–25, Oct. 2021, doi: 10.35290/RO.V2N3.2021.494.
- [6] Meltwater and We Are Social, "Digital 2023 April Global Statshot Report — DataReportal – Global Digital Insights." Accessed: Jun. 29, 2023. [Online]. Available: <https://datareportal.com/reports/digital-2023-april-global-statshot>
- [7] Y. Liu, "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *J Mark*, vol. 70, no. 3, pp. 74–89, Oct. 2018, doi: 10.1509/JMKG.70.3.074.
- [8] A. Constantin and R. Robert, "Addressing Sentiment Analysis Challenges within AI Media Platform: The Enabling Role of an AI Powered Chatbot," *Risk in Contemporary Economy*, pp. 399–406, 2021, Accessed: Mar. 22, 2023. [Online]. Available: <https://ideas.repec.org/a/ddj/fserec/y2021p399-406.html>
- [9] K. Shah, H. Patel, D. Sanghvi, and M. Shah, "A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text

Classification,” *Augmented Human Research 2020 5:1*, vol. 5, no. 1, pp. 1–16, Mar. 2020, doi: 10.1007/S41133-020-00032-0.

- [10] S. Chirgaiya, D. Sukheja, N. Shrivastava, and R. Rawat, “Analysis of sentiment based movie reviews using machine learning techniques,” *Journal of Intelligent & Fuzzy Systems*, vol. 41, no. 5, pp. 5449–5456, Jan. 2021, doi: 10.3233/JIFS-189866.
- [11] P. Sánchez-Holgado, M. Martín-Merino Acera, and D. Blanco Herrero, “Del data-driven al data-feeling: análisis de sentimiento en tiempo real de mensajes en español sobre divulgación científica usando técnicas de aprendizaje automático,” *Anuario Electrónico de Estudios en Comunicación Social “Disertaciones,”* vol. 13, no. 1, Jan. 2020, doi: 10.12804/REVISTAS.UROSARIO.EDU.CO/DISERTACIONES/A.7691.
- [12] K. Rodriguez-Diaz and Y. Haber-Guerra, “Sentiment analysis on Twitter applied to Donald Trump’s #impeachment,” *Revista Mediterranea de Comunicacion*, vol. 11, no. 2, pp. 199–213, Jul. 2020, doi: 10.14198/MEDCOM2020.11.2.23.
- [13] K. Dashtipour, M. Gogate, A. Adeel, H. Larjani, and A. Hussain, “Sentiment Analysis of Persian Movie Reviews Using Deep Learning,” *Entropy 2021, Vol. 23, Page 596*, vol. 23, no. 5, p. 596, May 2021, doi: 10.3390/E23050596.
- [14] G. Rosenbrock, S. Trossero, P. Goette, M. E. Llorente, A. Pascal, and D. Cían, “Técnicas de análisis de sentimientos aplicadas a la extracción de opiniones en el lenguaje español,” *XXI Workshop de Investigadores en Ciencias de la Computación*, pp. 971–974, 2019, Accessed: Aug. 14, 2023. [Online]. Available: <http://sedici.unlp.edu.ar/handle/10915/77029>
- [15] J. Aguilera, D. I. H. Farías, R. M. Ortega-Mendoza, and M. Montes-y-Gómez, “Depression and anorexia detection in social media as a one-class classification problem,” *Applied Intelligence*, vol. 51, no. 8, pp. 6088–6103, Aug. 2021, doi: 10.1007/S10489-020-02131-2/METRICS.
- [16] C. A. Arango Pastrana and C. F. Osorio Andrade, “Aislamiento social obligatorio: un análisis de sentimientos mediante machine learning,” *Suma de Negocios*, vol. 12, no. 26, pp. 1–13, Dec. 2021, doi: 10.14349/SUMNEG/2021.V12.N26.A1.

- [17] N. Öztürk and S. Ayvaz, "Sentiment analysis on Twitter: A text mining approach to the Syrian refugee crisis," *Telematics and Informatics*, vol. 35, no. 1, pp. 136–147, Apr. 2018, doi: 10.1016/J.TELE.2017.10.006.
- [18] M. A. Pérez Abelleira and C. A. Cardoso, "Minería de texto para la categorización automática de documentos," *Uniwersytet śląski*, vol. 7, no. 1, pp. 343–354, 2010, doi: 10.2/JQUERY.MIN.JS.
- [19] M. Contreras Barrera, "Text mining: a current view una visión actual," *Biblioteca Universitaria*, vol. 17, no. 2, pp. 129–138, 2014, Accessed: Oct. 23, 2022. [Online]. Available: <http://www.emc.com/leadership/programs/digital-universe.htm>
- [20] V. Gupta and G. S. Lehal Professor, "A Survey of Text Mining Techniques and Applications," *Journal of emerging technologies in web intelligence*, vol. 1, no. 2009, pp. 60–76, Aug. 2009, Accessed: Oct. 23, 2022. [Online]. Available: <http://www.jetwi.us/uploadfile/2014/1230/20141230112729939.pdf>
- [21] R. Sánchez del Hoyo, "Análisis de sentimientos con Twitter: turismo y política electoral," Universidad de Sevilla, Sevilla, 2019. Accessed: Oct. 23, 2022. [Online]. Available: <https://idus.us.es/handle/11441/90023>
- [22] B. Pang and L. Lee, "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts," *42nd ACL*, pp. 1–8, Sep. 2004, doi: 10.48550/arxiv.cs/0409058.
- [23] S. Sharma and G. Dutta, "A Novel Approach for Creating Sentiment Lexicon for Movie Reviews using Star Ratings," *International Journal on Emerging Technologies*, vol. 11, no. 1, pp. 376–386, 2020, Accessed: Jan. 29, 2023. [Online]. Available: [www.researchtrend.net](http://www.researchtrend.net)
- [24] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment Classification using Machine Learning Techniques," *Empirical Methods in Natural Language Processing*, pp. 79–86, 2002, doi: 10.3115/1118693.1118704.
- [25] H. Rui, Y. Liu, and A. Whinston, "Whose and what chatter matters? The effect of tweets on movie sales," *Decis Support Syst*, vol. 55, no. 4, pp. 863–870, Nov. 2013, doi: 10.1016/J.DSS.2012.12.022.
- [26] M.-T. Martín-Valdivia, E. Martínez-Cámara, J.-M. Perea-Ortega, and L. A. Ureña-López, "Sentiment polarity detection in Spanish reviews combining



- supervised and unsupervised approaches,” *Expert Syst Appl*, vol. 40, no. 10, pp. 3934–3942, Aug. 2013, doi: 10.1016/j.eswa.2012.12.084.
- [27] M. So Kim, J. Woo Kim, and C. Jing, “Comparison of Domain-Specific Lexicon Construction Methods for Sentiment Analysis,” *Advanced Science and Technology Letters*, pp. 1–5, 2016, doi: 10.14257/astl.2016.135.38.
- [28] B. Ay Karakuş, M. Talo, İ. R. Hallaç, and G. Aydin, “Evaluating deep learning models for sentiment classification,” *Concurr Comput*, vol. 30, no. 21, p. e4783, Nov. 2018, doi: 10.1002/CPE.4783.
- [29] W. Duan, B. Gu, and A. B. Whinston, “The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry,” *Journal of Retailing*, vol. 84, no. 2, pp. 233–242, Jun. 2008, doi: 10.1016/J.JRETAI.2008.04.005.
- [30] K. Kumar, B. S. Harish, and H. K. Darshan, “Sentiment Analysis on IMDb Movie Reviews Using Hybrid Feature Extraction Method,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 5, no. 5, p. 109, 2019, doi: 10.9781/IJIMAI.2018.12.005.
- [31] A. Timmaraju and V. Khanna, “Sentiment Analysis on Movie Reviews using Recursive and Recurrent Neural Network Architectures,” *Computer Science*, pp. 1–6, 2015.
- [32] N. M. Ali, M. M. Abd El Hamid, and A. Youssif, “Sentiment Analysis for Movies Reviews Dataset Using Deep Learning Models,” *International Journal of Data Mining & Knowledge Management Process*, vol. 9, no. 2/3, Jun. 14, 2019. Accessed: Jan. 29, 2023. [Online]. Available: <https://papers.ssrn.com/abstract=3403985>
- [33] A. S. Manek, P. D. Shenoy, M. C. Mohan, and K. R. Venugopal, “Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier,” *World Wide Web*, vol. 20, no. 2, pp. 135–154, Mar. 2017, doi: 10.1007/S11280-015-0381-X/TABLES/9.
- [34] A. Humberto Chaviano, “Técnicas de aprendizaje supervisado y no supervisado para el aprendizaje automatizado de computadoras,” *Memorias del primer Congreso Internacional de Ciencias Pedagógicas: Por una educación integral, participativa e incluyente, 2015*, ISBN 978-9942-17-011-8, págs. 549–564, pp. 549–564, 2015, Accessed: Aug. 14, 2023. [Online]. Available:

<https://dialnet.unirioja.es/servlet/articulo?codigo=7192675&info=resumen&idioma=SPA>

- [35] S. J. Russell and P. Norvig, *Inteligencia artificial: un enfoque moderno (2a. ed.)*. 2013. Accessed: Sep. 13, 2023. [Online]. Available: <https://bv.unir.net:3555/es/ereader/unir/45310>
- [36] D. Y. Ramírez Mendoza, "Métodos de machine learning con algoritmos de clúster no supervisados, una alternativa de segmentación de las pymes colombianas para plantear estrategias de acuerdo con sus condiciones económicas," 2022. Accessed: Sep. 11, 2023. [Online]. Available: <http://repository.eafit.edu.co/handle/10784/31580>
- [37] R. Michalski, "Understanding the Nature of Learning: Issues and Research Directions," 1985, Accessed: Sep. 13, 2023. [Online]. Available: <http://ebot.gmu.edu/bitstream/handle/1920/1585/85-09.pdf?sequence=1&isAllowed=y>
- [38] H. Liu and B. Lang, "Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey," *Applied Sciences* 2019, Vol. 9, Page 4396, vol. 9, no. 20, p. 4396, Oct. 2019, doi: 10.3390/APP9204396.
- [39] V. R. Konasani and S. Kadre, *INTRODUCTION TO MACHINE LEARNING AND DEEP LEARNING*. McGraw-Hill Education, 2021. Accessed: Nov. 10, 2023. [Online]. Available: <https://www-accessengineeringlibrary-com.ugto.idm.oclc.org/content/book/9781260462296/chapter/chapter1>
- [40] L. J. Sandoval Serrano, "Algoritmos de aprendizaje automático para análisis y predicción de datos," vol. 11, Oct. 2018, Accessed: Jan. 21, 2024. [Online]. Available: <http://redicces.org.sv/jspui/handle/10972/3626>
- [41] B. B. Martínez, "Minería de datos," *Cómo hallar una aguja en un pajar. Ingenierías*, vol. 14, pp. 53–66, 2001, Accessed: Jan. 21, 2024. [Online]. Available: <https://www.cs.buap.mx/~bbeltran/NotasMD.pdf>
- [42] C. P. López and D. S. González, "Minería de datos. Técnicas y herramientas: técnicas y herramientas," 2007, Accessed: Jan. 21, 2024. [Online]. Available: [https://books.google.es/books?hl=es&lr=&id=wz-D\\_8uPFCEC&oi=fnd&pg=PR4&dq=miner%C3%ADa+de+datos&ots=TkT3n0t8H&sig=v4e-2Fleq8hn8mK1VpF7lbtvn0](https://books.google.es/books?hl=es&lr=&id=wz-D_8uPFCEC&oi=fnd&pg=PR4&dq=miner%C3%ADa+de+datos&ots=TkT3n0t8H&sig=v4e-2Fleq8hn8mK1VpF7lbtvn0)

- [43] B. Liu, *Sentiment Analysis and Opinion Mining*, Springer Nature. 2022. Accessed: Nov. 10, 2023. [Online]. Available: [https://books.google.es/books?id=xYhyEAAAQBAJ&dq=opini%C3%B3n+mining&lr=&hl=es&source=gbs\\_navlinks\\_s](https://books.google.es/books?id=xYhyEAAAQBAJ&dq=opini%C3%B3n+mining&lr=&hl=es&source=gbs_navlinks_s)
- [44] S. Bhatia, P. Chaudhary, and N. Dey, *Opinion Mining in Information Retrieval*. in SpringerBriefs in Applied Sciences and Technology. Singapore: Springer Singapore, 2020. doi: 10.1007/978-981-15-5043-0.
- [45] M. Augusto Cortez Vásquez, M. Hugo Vega Huerta, L. Jaime, and P. Quispe, "Procesamiento de lenguaje natural," *Revista de investigación de Sistemas e Informática*, vol. 6, pp. 45–54, 2009, Accessed: Nov. 10, 2023. [Online]. Available: <https://revistasinvestigacion.unmsm.edu.pe/index.php/sistem/article/view/5923>
- [46] A. Pantoja Chaves, "Los nuevos medios de comunicación social: las redes sociales," 2011, Accessed: Jan. 17, 2024. [Online]. Available: <http://148.202.167.116:8080/xmlui/handle/123456789/3535>
- [47] C. Salazar, J. Aguilar, J. Monsalve-Pulido, and E. Montoya, "Análisis de Sentimientos/Polaridad en diferentes tipos de documentos," *Iberian Journal of Information Systems and Technologies*, 2020.
- [48] C. Musto, G. Semeraro, and M. Polignano, "A comparison of Lexicon-based approaches for Sentiment Analysis of microblog posts," *Proceedings of the 8th International Workshop on Information Filtering and Retrieval*, pp. 66–68, 2014, Accessed: Jan. 29, 2023. [Online]. Available: <http://aiia2014.di.unipi.it/dart/>
- [49] M. Ramos, "Hacia un nuevo recurso léxico: ¿ fusión entre corpus y diccionario?," *A survey of corpus-based research*, pp. 1–17, 2009, Accessed: Jan. 17, 2024. [Online]. Available: <http://www.um.es/lacell/aelinco/contenido/pdf/82.pdf>
- [50] S. Baccianella, A. Esuli, and F. Sebastiani, "SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining," *In Lrec*, vol. 10, pp. 2200–2204, Apr. 2010, Accessed: Jan. 29, 2023. [Online]. Available: <http://wordnetcode.princeton>.

- [51] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-Based Methods for Sentiment Analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, Jun. 2011, doi: 10.1162/COLI\_A\_00049.
- [52] J. Pennebaker, M. Francis, and R. Booth, "Linguistic inquiry and word count (LIWC)," *Erlbaum Publishers*, pp. 1–24, Jan. 1999, Accessed: Jan. 29, 2023. [Online]. Available: [https://www.researchgate.net/publication/246699633\\_Linguistic\\_inquiry\\_and\\_word\\_count\\_LIWC](https://www.researchgate.net/publication/246699633_Linguistic_inquiry_and_word_count_LIWC)
- [53] M. Thelwall, "The Heart and Soul of the Web? Sentiment Strength Detection in the Social Web with SentiStrength," *Cyberemotions*, pp. 119–134, 2017, doi: 10.1007/978-3-319-43639-5\_7.
- [54] L. Gatti, M. Guerini, and M. Turchi, "SentiWords: Deriving a High Precision and High Coverage Lexicon for Sentiment Analysis," *IEEE Trans Affect Comput*, vol. 7, no. 4, pp. 409–421, 2016, Accessed: Jan. 29, 2023. [Online]. Available: <https://arxiv.org/pdf/1510.09079.pdf>
- [55] M. M. Bradley and P. J. Lang, "Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings," *Technical Report C-1*, pp. 1–49, 1999, Accessed: Jan. 29, 2023. [Online]. Available: <https://pdodds.w3.uvm.edu/teaching/courses/2009-08UVM-300/docs/others/everything/bradley1999a.pdf>
- [56] P. J. Stone, D. C. Dunphy, and M. S. Smith, "The general inquirer: A computer approach to content analysis. - PsycNET," *M.I.T. Press*, pp. 651–660, 1966, Accessed: Jan. 29, 2023. [Online]. Available: <https://psycnet.apa.org/record/1967-04539-000>
- [57] P. D. Turney, "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews," *Empirical Methods in Natural Language Processing*, pp. 1–8, Dec. 2002, doi: 10.48550/arxiv.cs/0212032.
- [58] P. D. Turney and M. L. Littman, "Measuring praise and criticism," *ACM Transactions on Information Systems (TOIS)*, vol. 21, no. 4, pp. 315–346, Oct. 2003, doi: 10.1145/944012.944013.
- [59] V. Hatzivassiloglou and K. R. McKeown, "Predicting the Semantic Orientation of Adjectives," *Association for Computational Linguistics*, pp. 174–181, 1997, doi: 10.3115/976909.979640.

- [60] C. A. Alvarez Vazquez, "Exploración para la identificación automática de palabras con polaridad," 2020. Accessed: Oct. 31, 2023. [Online]. Available: <https://rinacional.tecnm.mx/jspui/handle/TecNM/4165>
- [61] B. Liu and L. Zhang, "A survey of opinion mining and sentiment analysis," *Mining Text Data*, vol. 9781461432234, pp. 415–463, Aug. 2012, doi: 10.1007/978-1-4614-3223-4\_13/COVER.
- [62] J. C. Alcalde, D. Isern, A. Carles, and V. Royo, "Análisis de sentimiento de textos basado en opiniones de películas usando algoritmos de aprendizaje computacional," Jun. 2021, Accessed: Oct. 30, 2023. [Online]. Available: <https://openaccess.uoc.edu/handle/10609/132328>
- [63] Y. D. Aguirre Ascona, "Métodos de aprendizaje supervisado para la predicción de diabetes: una revisión sistemática de la literatura," *Universidad Peruana Unión*, Dec. 2019, Accessed: Jan. 15, 2024. [Online]. Available: <https://repositorio.upeu.edu.pe/handle/20.500.12840/2511>
- [64] M. Salina, J. Osio, M. Cappelletti, and M. Morales, "Técnicas de análisis de sentimientos aplicadas a la valoración de opiniones en el lenguaje español," *Memorias del Congreso Argentino en Ciencias de la Computación - CACIC 2021*, vol. 1, no. 1, pp. 31–40, 2021, Accessed: Jul. 03, 2023. [Online]. Available: <http://sedici.unlp.edu.ar/handle/10915/130344>
- [65] N. Salmi and Z. Rustam, "Naïve Bayes Classifier Models for Predicting the Colon Cancer," *IOP Conf Ser Mater Sci Eng*, vol. 546, no. 5, p. 052068, Jun. 2019, doi: 10.1088/1757-899X/546/5/052068.
- [66] F. I. Adiba, T. Islam, M. S. Kaiser, M. Mahmud, and M. A. Rahman, "Effect of Corpora on Classification of Fake News using Naive Bayes Classifier," *International Journal of Automation, Artificial Intelligence and Machine Learning*, vol. 1, no. 1, pp. 80–92, Oct. 2020, Accessed: Jul. 05, 2023. [Online]. Available: <https://researchlakejournals.com/index.php/AAIML/article/view/45>
- [67] C. Beltrán and I. Barbona, "Regresión Logística y Árboles de Clasificación. Un estudio de simulación para la comparación en el caso de grupos balanceados y desbalanceados.," 2019, Accessed: Jul. 02, 2023. [Online]. Available: <http://rephip.unr.edu.ar/xmlui/handle/2133/14285>
- [68] R. Chaves, B. Guido, and V. Chavarría, "Clasificación de tumores mamarios utilizando regresión logística y KNN," *Revista Serengeti*, vol. 4, pp. 1–10, 2021, Accessed: Jul. 02, 2023. [Online]. Available:

<https://serengueti.fce.ucr.ac.cr/images/2023/03/16/Articulos/Estudio%20clasificacion%20tumores%20mamarios.pdf>

- [69] E. A. Galindo, J. A. Perdomo, and J. C. Figueroa-García, “Estudio comparativo entre máquinas de soporte vectorial multiclase, redes neuronales artificiales y sistema de inferencia neuro-difuso auto organizado para problemas de clasificación,” *Información tecnológica*, vol. 31, no. 1, pp. 273–286, Feb. 2020, doi: 10.4067/S0718-07642020000100273.
- [70] K. Shah, H. Patel, D. Sanghvi, and M. Shah, “A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification,” *Augmented Human Research 2020 5:1*, vol. 5, no. 1, pp. 1–16, Mar. 2020, doi: 10.1007/S41133-020-00032-0.
- [71] R. Devika, S. V. Avilala, and V. Subramaniaswamy, “Comparative study of classifier for chronic kidney disease prediction using naive bayes, KNN and random forest,” *Proceedings of the 3rd International Conference on Computing Methodologies and Communication, ICCMC 2019*, pp. 679–684, Mar. 2019, doi: 10.1109/ICCMC.2019.8819654.
- [72] F. M. Javed Mehedi Shamrat *et al.*, “Sentiment analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm,” *Article in Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 1, pp. 463–470, 2021, doi: 10.11591/ijeecs.v23.i1.pp463-470.
- [73] J. L. Speiser, M. E. Miller, J. Tooze, and E. Ip, “A comparison of random forest variable selection methods for classification prediction modeling,” *Expert Syst Appl*, vol. 134, pp. 93–101, Nov. 2019, doi: 10.1016/J.ESWA.2019.05.028.
- [74] P. Karthika, R. Murugeswari, and R. Manoranjithem, “Sentiment Analysis of Social Media Network Using Random Forest Algorithm,” *IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing, INCOS 2019*, Apr. 2019, doi: 10.1109/INCOS45849.2019.8951367.
- [75] Y. Al Amrani, M. Lazaar, and K. E. El Kadirp, “Random Forest and Support Vector Machine based Hybrid Approach to Sentiment Analysis,” *Procedia Comput Sci*, vol. 127, pp. 511–520, Jan. 2018, doi: 10.1016/J.PROCS.2018.01.150.

- [76] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts, "Learning Word Vectors for Sentiment Analysis," *the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*, pp. 142–150, 2011.
- [77] M. Yasen and S. Tedmori, "Movies reviews sentiment analysis and classification," *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology, JEEIT 2019 - Proceedings*, pp. 860–865, May 2019, doi: 10.1109/JEEIT.2019.8717422.
- [78] P. A. Pauli, "Análisis de sentimiento: comparación de algoritmos predictivos y métodos utilizando un lexicon español," Instituto Tecnológico De Buenos Aires, Buenos Aires, 2019. Accessed: Jul. 05, 2023. [Online]. Available: <http://ri.itba.edu.ar/handle/123456789/1782>
- [79] J. Antonio. Gil Pascual, *Minería de texto con R*. UNED - Universidad Nacional de Educación a Distancia, 2021. Accessed: Aug. 08, 2023. [Online]. Available: [https://books.google.com/books/about/Miner%C3%ADa\\_de\\_texto\\_con\\_R.html?hl=es&id=g95DEAAAQBAJ](https://books.google.com/books/about/Miner%C3%ADa_de_texto_con_R.html?hl=es&id=g95DEAAAQBAJ)
- [80] C. Hernández-Ávila, N. E.- Alerta, R. científica del, and undefined 2019, "Introducción a los tipos de muestreo," *camjol.info*, 2019, doi: 10.5377/alerta.v2i1.7535.
- [81] C. L. Corso, "Aplicación de algoritmos de clasificación supervisada usando Weka," *Córdoba: Universidad Tecnológica Nacional*, 2009, Accessed: Nov. 14, 2023. [Online]. Available: [https://www.investigacion.frc.utn.edu.ar/labsis/Publicaciones/congresos\\_labsis/cynthia/CNIT\\_2009\\_Aplicacion\\_Algoritmos\\_Weka.pdf](https://www.investigacion.frc.utn.edu.ar/labsis/Publicaciones/congresos_labsis/cynthia/CNIT_2009_Aplicacion_Algoritmos_Weka.pdf)
- [82] E. O. Omuya, G. Okeyo, and M. Kimwele, "Sentiment analysis on social media tweets using dimensionality reduction and natural language processing," *Engineering Reports*, vol. 5, no. 3, Mar. 2023, doi: 10.1002/ENG2.12579.
- [83] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review 2022* 55:7, vol. 55, no. 7, pp. 5731–5780, Feb. 2022, doi: 10.1007/S10462-022-10144-1.
- [84] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artif Intell Rev*, vol. 53, no. 6, pp. 4335–4385, Aug. 2020, doi: 10.1007/S10462-019-09794-5/METRICS.





---

## Apéndice A

---

Lista de palabras de paro utilizado en el sistema, dando un total de 317 palabras:

- why
- both
- throughout
- they
- found
- me
- couldn't
- through
- each
- seemed
- already
- nine
- done
- which
- behind
- here
- in
- so
- has
- whatever
- full
- always
- them
- under
- whenever
- such
- its
- made
- their
- whole
- thereafter
- against
- interest
- into
- but
- what
- third
- he
- rather
- between
- latterly
- are
- fire
- become
- back
- us
- hundred
- another
- detail
- call
- however
- re
- together
- most
- may
- very
- should
- whereafter
- without
- these
- moreover
- whence
- seems
- also
- above
- per
- thin
- nobody
- off
- every
- top
- part
- ten
- to
- a
- it
- sometimes
- sometime
- de
- please
- whoever
- that
- whereby
- thick
- the
- hence
- amount
- out
- con
- an
- had
- some
- forty
- anywhere
- is
- thereupon

- bill
- other
- side
- because
- from
- beside
- during
- you
- hereby
- was
- we
- becomes
- eleven
- formerly
- herself
- too
- would
- amongst
- four
- except
- sincere
- can't
- around
- co
- due
- have
- and
- not
- whereas
- bottom
- put
- see
- toward
- your
- go
- show
- inc
- serious
- hereafter
- none
- must
- now
- were
- below
- about
- take
- hereupon
- wherever
- name
- only
- indeed
- somewhere
- themselves
- wherein
- meanwhile
- nowhere
- less
- well
- no one
- nothing
- move
- while
- via
- along
- one
- sixty
- mill
- on
- once
- being
- becoming
- anyone
- eight
- more
- of
- upon
- even
- beforehand
- ever
- might
- no
- last
- either
- his
- further
- this
- became
- anyhow
- twelve
- nevertheless
- enough
- everywhere
- yourselves
- down
- onto
- who
- beyond
- perhaps
- I
- ours
- others
- with
- there
- as
- therein
- almost
- over
- much

- several
- many
- hers
- whither
- fifty
- two
- give
- whether
- still
- someone
- her
- five
- whereupon
- find
- next
- itself
- thru
- fill
- somehow
- own
- be
- three
- everything
- she
- by
- otherwise
- how
- least
- empty
- since
- get
- yet
- nor
- herein
- i.e.
- anyway
- ltd
- could
- besides
- else
- before
- etc.
- then
- our
- neither
- my
- twenty
- fifteen
- can
- thereby
- whose
- latter
- all
- e.g.
- yours
- across
- do
- thence
- therefore
- than
- cry
- although
- thus
- will
- if
- when
- am
- often
- system
- anything
- hasn't
- cannot
- among
- something
- un
- yourself
- though
- seem
- ourselves
- until
- few
- afterwards
- whom
- within
- elsewhere
- first
- again
- after
- up
- for
- mostly
- same
- former
- amongst
- describe
- myself
- seeming
- been
- never
- those
- alone
- him
- himself
- namely
- keep
- mine
- six
- at

- or
- where
- front
- towards
- any
- everyone

## Apéndice B

Significado de abreviaturas: Exactitud (E), Precisión (P), Recall (R), Cantidad de Opiniones (CA), Punto de Partida (PP), Lematización (Le), Eliminación de acentos (EA), Eliminación Mayúsculas (EM), Eliminación de palabras de paro (EP).

Todos los resultados:

| <i>Metodo</i> | <i>E</i> | <i>P</i> | <i>R</i> | <i>F1-score</i> | <i>CO</i> | <i>Ciclo</i> | <i>Tipo</i> |
|---------------|----------|----------|----------|-----------------|-----------|--------------|-------------|
| NB            | 0.88     | 0.88     | 0.82     | 0.85            | 200       | 1            | PP          |
| RL            | 0.78     | 0.83     | 0.59     | 0.69            | 200       | 1            | PP          |
| K-VC          | 0.58     | 0.50     | 0.71     | 0.59            | 200       | 1            | PP          |
| BA            | 0.58     | 0.50     | 0.41     | 0.45            | 200       | 1            | PP          |
| MSV           | 0.45     | 0.42     | 0.76     | 0.54            | 200       | 1            | PP          |
| NB            | 0.98     | 0.98     | 0.98     | 0.98            | 400       | 2            | PP          |
| RL            | 0.89     | 0.93     | 0.87     | 0.90            | 400       | 2            | PP          |
| MSV           | 0.73     | 0.93     | 0.56     | 0.69            | 400       | 2            | PP          |
| BA            | 0.69     | 0.83     | 0.56     | 0.67            | 400       | 2            | PP          |
| K-VC          | 0.58     | 0.61     | 0.69     | 0.65            | 400       | 2            | PP          |
| NB            | 0.93     | 0.92     | 0.93     | 0.93            | 600       | 3            | PP          |
| RL            | 0.85     | 0.85     | 0.85     | 0.85            | 600       | 3            | PP          |
| BA            | 0.73     | 0.84     | 0.59     | 0.69            | 600       | 3            | PP          |
| MSV           | 0.64     | 0.69     | 0.54     | 0.61            | 600       | 3            | PP          |
| K-VC          | 0.58     | 0.58     | 0.64     | 0.61            | 600       | 3            | PP          |
| NB            | 0.91     | 0.94     | 0.89     | 0.91            | 800       | 4            | PP          |
| RL            | 0.89     | 0.92     | 0.88     | 0.90            | 800       | 4            | PP          |
| BA            | 0.73     | 0.80     | 0.69     | 0.74            | 800       | 4            | PP          |
| MSV           | 0.66     | 0.77     | 0.55     | 0.64            | 800       | 4            | PP          |
| K-VC          | 0.62     | 0.65     | 0.67     | 0.66            | 800       | 4            | PP          |
| NB            | 0.90     | 0.91     | 0.88     | 0.90            | 1000      | 5            | PP          |
| RL            | 0.81     | 0.80     | 0.85     | 0.82            | 1000      | 5            | PP          |
| MSV           | 0.72     | 0.74     | 0.69     | 0.72            | 1000      | 5            | PP          |
| BA            | 0.69     | 0.75     | 0.59     | 0.66            | 1000      | 5            | PP          |
| K-VC          | 0.61     | 0.61     | 0.68     | 0.65            | 1000      | 5            | PP          |
| NB            | 0.90     | 0.92     | 0.85     | 0.89            | 1200      | 6            | PP          |
| RL            | 0.81     | 0.80     | 0.81     | 0.80            | 1200      | 6            | PP          |
| MSV           | 0.73     | 0.70     | 0.76     | 0.73            | 1200      | 6            | PP          |
| BA            | 0.72     | 0.72     | 0.67     | 0.69            | 1200      | 6            | PP          |
| K-VC          | 0.59     | 0.56     | 0.68     | 0.61            | 1200      | 6            | PP          |
| NB            | 0.91     | 0.95     | 0.89     | 0.92            | 1400      | 7            | PP          |
| RL            | 0.88     | 0.89     | 0.88     | 0.88            | 1400      | 7            | PP          |
| MSV           | 0.71     | 0.76     | 0.71     | 0.73            | 1400      | 7            | PP          |
| BA            | 0.69     | 0.80     | 0.58     | 0.67            | 1400      | 7            | PP          |
| K-VC          | 0.60     | 0.61     | 0.72     | 0.66            | 1400      | 7            | PP          |

|      |      |      |      |      |      |    |    |
|------|------|------|------|------|------|----|----|
| NB   | 0.91 | 0.90 | 0.90 | 0.90 | 1600 | 8  | PP |
| RL   | 0.86 | 0.86 | 0.85 | 0.85 | 1600 | 8  | PP |
| MSV  | 0.71 | 0.67 | 0.78 | 0.72 | 1600 | 8  | PP |
| BA   | 0.71 | 0.72 | 0.66 | 0.68 | 1600 | 8  | PP |
| K-VC | 0.57 | 0.54 | 0.67 | 0.60 | 1600 | 8  | PP |
| NB   | 0.92 | 0.92 | 0.93 | 0.92 | 1800 | 9  | PP |
| RL   | 0.85 | 0.85 | 0.85 | 0.85 | 1800 | 9  | PP |
| MSV  | 0.74 | 0.73 | 0.80 | 0.76 | 1800 | 9  | PP |
| BA   | 0.69 | 0.74 | 0.60 | 0.66 | 1800 | 9  | PP |
| K-VC | 0.58 | 0.57 | 0.67 | 0.61 | 1800 | 9  | PP |
| NB   | 0.91 | 0.91 | 0.90 | 0.91 | 2000 | 10 | PP |
| RL   | 0.90 | 0.89 | 0.91 | 0.90 | 2000 | 10 | PP |
| MSV  | 0.77 | 0.75 | 0.80 | 0.78 | 2000 | 10 | PP |
| BA   | 0.74 | 0.76 | 0.69 | 0.72 | 2000 | 10 | PP |
| K-VC | 0.56 | 0.55 | 0.67 | 0.60 | 2000 | 10 | PP |
| NB   | 0.88 | 0.88 | 0.82 | 0.85 | 200  | 11 | EM |
| RL   | 0.80 | 0.85 | 0.65 | 0.73 | 200  | 11 | EM |
| BA   | 0.75 | 0.67 | 0.82 | 0.74 | 200  | 11 | EM |
| K-VC | 0.60 | 0.53 | 0.47 | 0.50 | 200  | 11 | EM |
| MSV  | 0.50 | 0.46 | 1.00 | 0.63 | 200  | 11 | EM |
| NB   | 0.96 | 0.98 | 0.96 | 0.97 | 400  | 12 | EM |
| RL   | 0.89 | 0.89 | 0.91 | 0.90 | 400  | 12 | EM |
| BA   | 0.74 | 0.85 | 0.64 | 0.73 | 400  | 12 | EM |
| MSV  | 0.69 | 0.88 | 0.51 | 0.65 | 400  | 12 | EM |
| K-VC | 0.51 | 0.57 | 0.53 | 0.55 | 400  | 12 | EM |
| NB   | 0.91 | 0.90 | 0.92 | 0.91 | 600  | 13 | EM |
| RL   | 0.89 | 0.88 | 0.92 | 0.90 | 600  | 13 | EM |
| BA   | 0.74 | 0.77 | 0.70 | 0.74 | 600  | 13 | EM |
| MSV  | 0.71 | 0.75 | 0.64 | 0.69 | 600  | 13 | EM |
| K-VC | 0.62 | 0.64 | 0.57 | 0.60 | 600  | 13 | EM |
| NB   | 0.91 | 0.94 | 0.90 | 0.92 | 800  | 14 | EM |
| RL   | 0.85 | 0.88 | 0.84 | 0.86 | 800  | 14 | EM |
| BA   | 0.73 | 0.87 | 0.60 | 0.71 | 800  | 14 | EM |
| MSV  | 0.70 | 0.82 | 0.60 | 0.69 | 800  | 14 | EM |
| K-VC | 0.59 | 0.64 | 0.60 | 0.62 | 800  | 14 | EM |
| NB   | 0.89 | 0.89 | 0.90 | 0.90 | 1000 | 15 | EM |
| RL   | 0.84 | 0.84 | 0.85 | 0.84 | 1000 | 15 | EM |
| MSV  | 0.76 | 0.76 | 0.78 | 0.77 | 1000 | 15 | EM |
| BA   | 0.73 | 0.80 | 0.63 | 0.71 | 1000 | 15 | EM |
| K-VC | 0.56 | 0.57 | 0.58 | 0.57 | 1000 | 15 | EM |
| NB   | 0.90 | 0.92 | 0.86 | 0.89 | 1200 | 16 | EM |
| RL   | 0.83 | 0.83 | 0.82 | 0.82 | 1200 | 16 | EM |
| MSV  | 0.74 | 0.69 | 0.81 | 0.74 | 1200 | 16 | EM |
| BA   | 0.70 | 0.72 | 0.61 | 0.66 | 1200 | 16 | EM |
| K-VC | 0.63 | 0.59 | 0.68 | 0.63 | 1200 | 16 | EM |

|             |      |      |      |      |      |    |    |
|-------------|------|------|------|------|------|----|----|
| <i>NB</i>   | 0.93 | 0.96 | 0.90 | 0.93 | 1400 | 17 | EM |
| <i>RL</i>   | 0.87 | 0.90 | 0.86 | 0.88 | 1400 | 17 | EM |
| <i>MSV</i>  | 0.75 | 0.80 | 0.73 | 0.76 | 1400 | 17 | EM |
| <i>BA</i>   | 0.69 | 0.80 | 0.56 | 0.66 | 1400 | 17 | EM |
| <i>K-VC</i> | 0.64 | 0.66 | 0.69 | 0.67 | 1400 | 17 | EM |
| <i>NB</i>   | 0.92 | 0.93 | 0.90 | 0.91 | 1600 | 18 | EM |
| <i>RL</i>   | 0.86 | 0.85 | 0.86 | 0.85 | 1600 | 18 | EM |
| <i>MSV</i>  | 0.76 | 0.72 | 0.81 | 0.76 | 1600 | 18 | EM |
| <i>BA</i>   | 0.71 | 0.74 | 0.61 | 0.67 | 1600 | 18 | EM |
| <i>K-VC</i> | 0.61 | 0.59 | 0.67 | 0.62 | 1600 | 18 | EM |
| <i>NB</i>   | 0.92 | 0.93 | 0.92 | 0.92 | 1800 | 19 | EM |
| <i>RL</i>   | 0.86 | 0.87 | 0.86 | 0.87 | 1800 | 19 | EM |
| <i>MSV</i>  | 0.78 | 0.76 | 0.81 | 0.79 | 1800 | 19 | EM |
| <i>BA</i>   | 0.74 | 0.77 | 0.69 | 0.73 | 1800 | 19 | EM |
| <i>K-VC</i> | 0.59 | 0.57 | 0.73 | 0.64 | 1800 | 19 | EM |
| <i>NB</i>   | 0.92 | 0.92 | 0.92 | 0.92 | 2000 | 20 | EM |
| <i>RL</i>   | 0.89 | 0.88 | 0.89 | 0.88 | 2000 | 20 | EM |
| <i>MSV</i>  | 0.81 | 0.79 | 0.83 | 0.81 | 2000 | 20 | EM |
| <i>BA</i>   | 0.73 | 0.76 | 0.65 | 0.70 | 2000 | 20 | EM |
| <i>K-VC</i> | 0.63 | 0.61 | 0.73 | 0.66 | 2000 | 20 | EM |
| <i>NB</i>   | 0.88 | 0.88 | 0.82 | 0.85 | 200  | 21 | EA |
| <i>RL</i>   | 0.80 | 0.85 | 0.65 | 0.73 | 200  | 21 | EA |
| <i>BA</i>   | 0.65 | 0.58 | 0.65 | 0.61 | 200  | 21 | EA |
| <i>K-VC</i> | 0.60 | 0.53 | 0.47 | 0.50 | 200  | 21 | EA |
| <i>MSV</i>  | 0.50 | 0.46 | 1.00 | 0.63 | 200  | 21 | EA |
| <i>NB</i>   | 0.96 | 0.98 | 0.96 | 0.97 | 400  | 22 | EA |
| <i>RL</i>   | 0.89 | 0.89 | 0.91 | 0.90 | 400  | 22 | EA |
| <i>BA</i>   | 0.69 | 0.79 | 0.60 | 0.68 | 400  | 22 | EA |
| <i>MSV</i>  | 0.69 | 0.88 | 0.51 | 0.65 | 400  | 22 | EA |
| <i>K-VC</i> | 0.51 | 0.57 | 0.53 | 0.55 | 400  | 22 | EA |
| <i>NB</i>   | 0.91 | 0.90 | 0.92 | 0.91 | 600  | 23 | EA |
| <i>RL</i>   | 0.89 | 0.88 | 0.92 | 0.90 | 600  | 23 | EA |
| <i>BA</i>   | 0.72 | 0.75 | 0.66 | 0.70 | 600  | 23 | EA |
| <i>MSV</i>  | 0.71 | 0.75 | 0.64 | 0.69 | 600  | 23 | EA |
| <i>K-VC</i> | 0.62 | 0.64 | 0.57 | 0.60 | 600  | 23 | EA |
| <i>NB</i>   | 0.92 | 0.94 | 0.91 | 0.93 | 800  | 24 | EA |
| <i>RL</i>   | 0.85 | 0.88 | 0.84 | 0.86 | 800  | 24 | EA |
| <i>MSV</i>  | 0.70 | 0.82 | 0.60 | 0.69 | 800  | 24 | EA |
| <i>BA</i>   | 0.67 | 0.82 | 0.52 | 0.63 | 800  | 24 | EA |
| <i>K-VC</i> | 0.59 | 0.64 | 0.60 | 0.62 | 800  | 24 | EA |
| <i>NB</i>   | 0.89 | 0.89 | 0.90 | 0.90 | 1000 | 25 | EA |
| <i>RL</i>   | 0.84 | 0.84 | 0.85 | 0.84 | 1000 | 25 | EA |
| <i>MSV</i>  | 0.76 | 0.76 | 0.78 | 0.77 | 1000 | 25 | EA |
| <i>BA</i>   | 0.69 | 0.77 | 0.58 | 0.66 | 1000 | 25 | EA |
| <i>K-VC</i> | 0.56 | 0.57 | 0.58 | 0.57 | 1000 | 25 | EA |

|      |      |      |      |      |      |    |    |
|------|------|------|------|------|------|----|----|
| NB   | 0.90 | 0.92 | 0.86 | 0.89 | 1200 | 26 | EA |
| RL   | 0.84 | 0.83 | 0.82 | 0.83 | 1200 | 26 | EA |
| MSV  | 0.74 | 0.69 | 0.81 | 0.74 | 1200 | 26 | EA |
| BA   | 0.74 | 0.77 | 0.63 | 0.70 | 1200 | 26 | EA |
| K-VC | 0.63 | 0.59 | 0.68 | 0.63 | 1200 | 26 | EA |
| NB   | 0.93 | 0.96 | 0.90 | 0.93 | 1400 | 27 | EA |
| RL   | 0.87 | 0.90 | 0.86 | 0.88 | 1400 | 27 | EA |
| MSV  | 0.75 | 0.80 | 0.73 | 0.76 | 1400 | 27 | EA |
| BA   | 0.68 | 0.80 | 0.56 | 0.66 | 1400 | 27 | EA |
| K-VC | 0.64 | 0.66 | 0.69 | 0.67 | 1400 | 27 | EA |
| NB   | 0.92 | 0.93 | 0.90 | 0.91 | 1600 | 28 | EA |
| RL   | 0.86 | 0.85 | 0.86 | 0.85 | 1600 | 28 | EA |
| MSV  | 0.76 | 0.72 | 0.81 | 0.76 | 1600 | 28 | EA |
| BA   | 0.73 | 0.74 | 0.69 | 0.72 | 1600 | 28 | EA |
| K-VC | 0.61 | 0.59 | 0.67 | 0.62 | 1600 | 28 | EA |
| NB   | 0.92 | 0.93 | 0.92 | 0.92 | 1800 | 29 | EA |
| RL   | 0.86 | 0.87 | 0.86 | 0.87 | 1800 | 29 | EA |
| MSV  | 0.78 | 0.76 | 0.81 | 0.79 | 1800 | 29 | EA |
| BA   | 0.75 | 0.77 | 0.72 | 0.75 | 1800 | 29 | EA |
| K-VC | 0.59 | 0.57 | 0.73 | 0.64 | 1800 | 29 | EA |
| NB   | 0.92 | 0.92 | 0.92 | 0.92 | 2000 | 30 | EA |
| RL   | 0.89 | 0.88 | 0.89 | 0.88 | 2000 | 30 | EA |
| MSV  | 0.81 | 0.79 | 0.83 | 0.81 | 2000 | 30 | EA |
| BA   | 0.73 | 0.78 | 0.65 | 0.71 | 2000 | 30 | EA |
| K-VC | 0.63 | 0.61 | 0.73 | 0.66 | 2000 | 30 | EA |
| RL   | 0.90 | 0.88 | 0.88 | 0.88 | 200  | 31 | EP |
| NB   | 0.88 | 0.88 | 0.82 | 0.85 | 200  | 31 | EP |
| BA   | 0.70 | 0.65 | 0.65 | 0.65 | 200  | 31 | EP |
| K-VC | 0.65 | 0.57 | 0.76 | 0.65 | 200  | 31 | EP |
| MSV  | 0.63 | 0.53 | 1.00 | 0.69 | 200  | 31 | EP |
| NB   | 0.95 | 0.96 | 0.96 | 0.96 | 400  | 32 | EP |
| RL   | 0.85 | 0.87 | 0.87 | 0.87 | 400  | 32 | EP |
| MSV  | 0.85 | 0.88 | 0.84 | 0.86 | 400  | 32 | EP |
| BA   | 0.83 | 0.97 | 0.71 | 0.82 | 400  | 32 | EP |
| K-VC | 0.56 | 0.62 | 0.58 | 0.60 | 400  | 32 | EP |
| NB   | 0.93 | 0.92 | 0.93 | 0.93 | 600  | 33 | EP |
| RL   | 0.91 | 0.89 | 0.93 | 0.91 | 600  | 33 | EP |
| MSV  | 0.83 | 0.84 | 0.80 | 0.82 | 600  | 33 | EP |
| BA   | 0.81 | 0.81 | 0.82 | 0.81 | 600  | 33 | EP |
| K-VC | 0.52 | 0.53 | 0.49 | 0.51 | 600  | 33 | EP |
| NB   | 0.93 | 0.94 | 0.93 | 0.94 | 800  | 34 | EP |
| RL   | 0.89 | 0.94 | 0.85 | 0.89 | 800  | 34 | EP |
| MSV  | 0.82 | 0.91 | 0.75 | 0.82 | 800  | 34 | EP |
| BA   | 0.76 | 0.89 | 0.65 | 0.75 | 800  | 34 | EP |
| K-VC | 0.53 | 0.58 | 0.52 | 0.55 | 800  | 34 | EP |



|             |      |      |      |      |      |    |         |
|-------------|------|------|------|------|------|----|---------|
| <i>NB</i>   | 0.92 | 0.92 | 0.93 | 0.92 | 1000 | 35 | EP      |
| <i>RL</i>   | 0.89 | 0.86 | 0.92 | 0.89 | 1000 | 35 | EP      |
| <i>MSV</i>  | 0.81 | 0.78 | 0.88 | 0.83 | 1000 | 35 | EP      |
| <i>BA</i>   | 0.72 | 0.78 | 0.63 | 0.70 | 1000 | 35 | EP      |
| <i>K-VC</i> | 0.56 | 0.61 | 0.40 | 0.49 | 1000 | 35 | EP      |
| <i>NB</i>   | 0.91 | 0.93 | 0.89 | 0.91 | 1200 | 36 | EP      |
| <i>RL</i>   | 0.88 | 0.87 | 0.89 | 0.88 | 1200 | 36 | EP      |
| <i>MSV</i>  | 0.83 | 0.79 | 0.88 | 0.83 | 1200 | 36 | EP      |
| <i>BA</i>   | 0.70 | 0.73 | 0.60 | 0.66 | 1200 | 36 | EP      |
| <i>K-VC</i> | 0.63 | 0.61 | 0.60 | 0.60 | 1200 | 36 | EP      |
| <i>NB</i>   | 0.93 | 0.95 | 0.93 | 0.94 | 1400 | 37 | EP      |
| <i>RL</i>   | 0.90 | 0.91 | 0.91 | 0.91 | 1400 | 37 | EP      |
| <i>MSV</i>  | 0.86 | 0.90 | 0.84 | 0.87 | 1400 | 37 | EP      |
| <i>BA</i>   | 0.78 | 0.85 | 0.73 | 0.79 | 1400 | 37 | EP      |
| <i>K-VC</i> | 0.60 | 0.73 | 0.41 | 0.53 | 1400 | 37 | EP      |
| <i>NB</i>   | 0.92 | 0.92 | 0.91 | 0.92 | 1600 | 38 | EP      |
| <i>RL</i>   | 0.89 | 0.86 | 0.91 | 0.89 | 1600 | 38 | EP      |
| <i>MSV</i>  | 0.85 | 0.82 | 0.88 | 0.85 | 1600 | 38 | EP      |
| <i>BA</i>   | 0.77 | 0.80 | 0.69 | 0.74 | 1600 | 38 | EP      |
| <i>K-VC</i> | 0.65 | 0.68 | 0.53 | 0.59 | 1600 | 38 | EP      |
| <i>NB</i>   | 0.92 | 0.93 | 0.92 | 0.92 | 1800 | 39 | EP      |
| <i>RL</i>   | 0.88 | 0.86 | 0.91 | 0.89 | 1800 | 39 | EP      |
| <i>MSV</i>  | 0.85 | 0.83 | 0.89 | 0.86 | 1800 | 39 | EP      |
| <i>BA</i>   | 0.76 | 0.77 | 0.75 | 0.76 | 1800 | 39 | EP      |
| <i>K-VC</i> | 0.61 | 0.60 | 0.70 | 0.65 | 1800 | 39 | EP      |
| <i>NB</i>   | 0.92 | 0.91 | 0.92 | 0.92 | 2000 | 40 | EP      |
| <i>RL</i>   | 0.91 | 0.89 | 0.92 | 0.91 | 2000 | 40 | EP      |
| <i>MSV</i>  | 0.89 | 0.87 | 0.91 | 0.89 | 2000 | 40 | EP      |
| <i>BA</i>   | 0.75 | 0.75 | 0.75 | 0.75 | 2000 | 40 | EP      |
| <i>K-VC</i> | 0.64 | 0.61 | 0.74 | 0.67 | 2000 | 40 | EP      |
| <i>RL</i>   | 0.93 | 0.94 | 0.88 | 0.91 | 200  | 41 | EA y EP |
| <i>NB</i>   | 0.88 | 0.88 | 0.82 | 0.85 | 200  | 41 | EA y EP |
| <i>BA</i>   | 0.83 | 0.73 | 0.94 | 0.82 | 200  | 41 | EA y EP |
| <i>K-VC</i> | 0.65 | 0.57 | 0.76 | 0.65 | 200  | 41 | EA y EP |
| <i>MSV</i>  | 0.63 | 0.53 | 1.00 | 0.69 | 200  | 41 | EA y EP |
| <i>NB</i>   | 0.95 | 0.96 | 0.96 | 0.96 | 400  | 42 | EA y EP |
| <i>RL</i>   | 0.86 | 0.89 | 0.87 | 0.88 | 400  | 42 | EA y EP |
| <i>MSV</i>  | 0.85 | 0.88 | 0.84 | 0.86 | 400  | 42 | EA y EP |
| <i>BA</i>   | 0.79 | 0.85 | 0.76 | 0.80 | 400  | 42 | EA y EP |
| <i>K-VC</i> | 0.56 | 0.62 | 0.58 | 0.60 | 400  | 42 | EA y EP |
| <i>NB</i>   | 0.93 | 0.92 | 0.93 | 0.93 | 600  | 43 | EA y EP |
| <i>RL</i>   | 0.91 | 0.89 | 0.93 | 0.91 | 600  | 43 | EA y EP |
| <i>MSV</i>  | 0.83 | 0.84 | 0.80 | 0.82 | 600  | 43 | EA y EP |
| <i>BA</i>   | 0.77 | 0.82 | 0.69 | 0.75 | 600  | 43 | EA y EP |
| <i>K-VC</i> | 0.52 | 0.53 | 0.49 | 0.51 | 600  | 43 | EA y EP |

|      |      |      |      |      |      |    |         |
|------|------|------|------|------|------|----|---------|
| NB   | 0.93 | 0.94 | 0.93 | 0.94 | 800  | 44 | EA y EP |
| RL   | 0.89 | 0.94 | 0.85 | 0.89 | 800  | 44 | EA y EP |
| MSV  | 0.82 | 0.91 | 0.75 | 0.82 | 800  | 44 | EA y EP |
| BA   | 0.76 | 0.84 | 0.70 | 0.76 | 800  | 44 | EA y EP |
| K-VC | 0.53 | 0.58 | 0.52 | 0.55 | 800  | 44 | EA y EP |
| NB   | 0.92 | 0.92 | 0.93 | 0.92 | 1000 | 45 | EA y EP |
| RL   | 0.89 | 0.86 | 0.92 | 0.89 | 1000 | 45 | EA y EP |
| MSV  | 0.81 | 0.78 | 0.88 | 0.83 | 1000 | 45 | EA y EP |
| BA   | 0.76 | 0.80 | 0.72 | 0.76 | 1000 | 45 | EA y EP |
| K-VC | 0.56 | 0.61 | 0.40 | 0.49 | 1000 | 45 | EA y EP |
| NB   | 0.91 | 0.93 | 0.89 | 0.91 | 1200 | 46 | EA y EP |
| RL   | 0.88 | 0.87 | 0.89 | 0.88 | 1200 | 46 | EA y EP |
| MSV  | 0.83 | 0.79 | 0.88 | 0.83 | 1200 | 46 | EA y EP |
| BA   | 0.75 | 0.77 | 0.67 | 0.71 | 1200 | 46 | EA y EP |
| K-VC | 0.63 | 0.61 | 0.60 | 0.60 | 1200 | 46 | EA y EP |
| NB   | 0.93 | 0.95 | 0.93 | 0.94 | 1400 | 47 | EA y EP |
| RL   | 0.90 | 0.91 | 0.91 | 0.91 | 1400 | 47 | EA y EP |
| MSV  | 0.86 | 0.90 | 0.84 | 0.87 | 1400 | 47 | EA y EP |
| BA   | 0.77 | 0.86 | 0.69 | 0.76 | 1400 | 47 | EA y EP |
| K-VC | 0.60 | 0.73 | 0.41 | 0.53 | 1400 | 47 | EA y EP |
| NB   | 0.92 | 0.92 | 0.91 | 0.92 | 1600 | 48 | EA y EP |
| RL   | 0.89 | 0.86 | 0.91 | 0.89 | 1600 | 48 | EA y EP |
| MSV  | 0.85 | 0.82 | 0.88 | 0.85 | 1600 | 48 | EA y EP |
| BA   | 0.76 | 0.80 | 0.66 | 0.72 | 1600 | 48 | EA y EP |
| K-VC | 0.65 | 0.68 | 0.53 | 0.59 | 1600 | 48 | EA y EP |
| NB   | 0.92 | 0.93 | 0.92 | 0.92 | 1800 | 49 | EA y EP |
| RL   | 0.88 | 0.87 | 0.91 | 0.89 | 1800 | 49 | EA y EP |
| MSV  | 0.85 | 0.83 | 0.89 | 0.86 | 1800 | 49 | EA y EP |
| BA   | 0.74 | 0.79 | 0.66 | 0.72 | 1800 | 49 | EA y EP |
| K-VC | 0.61 | 0.60 | 0.70 | 0.65 | 1800 | 49 | EA y EP |
| NB   | 0.92 | 0.91 | 0.92 | 0.92 | 2000 | 50 | EA y EP |
| RL   | 0.91 | 0.89 | 0.92 | 0.91 | 2000 | 50 | EA y EP |
| MSV  | 0.89 | 0.87 | 0.91 | 0.89 | 2000 | 50 | EA y EP |
| BA   | 0.76 | 0.81 | 0.68 | 0.74 | 2000 | 50 | EA y EP |
| K-VC | 0.64 | 0.61 | 0.74 | 0.67 | 2000 | 50 | EA y EP |
| NB   | 0.88 | 0.88 | 0.82 | 0.85 | 200  | 51 | EA y EM |
| RL   | 0.80 | 0.85 | 0.65 | 0.73 | 200  | 51 | EA y EM |
| BA   | 0.65 | 0.58 | 0.65 | 0.61 | 200  | 51 | EA y EM |
| K-VC | 0.60 | 0.53 | 0.47 | 0.50 | 200  | 51 | EA y EM |
| MSV  | 0.50 | 0.46 | 1.00 | 0.63 | 200  | 51 | EA y EM |
| NB   | 0.96 | 0.98 | 0.96 | 0.97 | 400  | 52 | EA y EM |
| RL   | 0.89 | 0.89 | 0.91 | 0.90 | 400  | 52 | EA y EM |
| BA   | 0.69 | 0.79 | 0.60 | 0.68 | 400  | 52 | EA y EM |
| MSV  | 0.69 | 0.88 | 0.51 | 0.65 | 400  | 52 | EA y EM |
| K-VC | 0.51 | 0.57 | 0.53 | 0.55 | 400  | 52 | EA y EM |

|             |      |      |      |      |      |    |         |
|-------------|------|------|------|------|------|----|---------|
| <i>NB</i>   | 0.91 | 0.90 | 0.92 | 0.91 | 600  | 53 | EA y EM |
| <i>RL</i>   | 0.89 | 0.88 | 0.92 | 0.90 | 600  | 53 | EA y EM |
| <i>BA</i>   | 0.72 | 0.75 | 0.66 | 0.70 | 600  | 53 | EA y EM |
| <i>MSV</i>  | 0.71 | 0.75 | 0.64 | 0.69 | 600  | 53 | EA y EM |
| <i>K-VC</i> | 0.62 | 0.64 | 0.57 | 0.60 | 600  | 53 | EA y EM |
| <i>NB</i>   | 0.92 | 0.94 | 0.91 | 0.93 | 800  | 54 | EA y EM |
| <i>RL</i>   | 0.85 | 0.88 | 0.84 | 0.86 | 800  | 54 | EA y EM |
| <i>MSV</i>  | 0.70 | 0.82 | 0.60 | 0.69 | 800  | 54 | EA y EM |
| <i>BA</i>   | 0.67 | 0.82 | 0.52 | 0.63 | 800  | 54 | EA y EM |
| <i>K-VC</i> | 0.59 | 0.64 | 0.60 | 0.62 | 800  | 54 | EA y EM |
| <i>NB</i>   | 0.89 | 0.89 | 0.90 | 0.90 | 1000 | 55 | EA y EM |
| <i>RL</i>   | 0.84 | 0.84 | 0.85 | 0.84 | 1000 | 55 | EA y EM |
| <i>MSV</i>  | 0.76 | 0.76 | 0.78 | 0.77 | 1000 | 55 | EA y EM |
| <i>BA</i>   | 0.69 | 0.77 | 0.58 | 0.66 | 1000 | 55 | EA y EM |
| <i>K-VC</i> | 0.56 | 0.57 | 0.58 | 0.57 | 1000 | 55 | EA y EM |
| <i>NB</i>   | 0.90 | 0.92 | 0.86 | 0.89 | 1200 | 56 | EA y EM |
| <i>RL</i>   | 0.84 | 0.83 | 0.82 | 0.83 | 1200 | 56 | EA y EM |
| <i>MSV</i>  | 0.74 | 0.69 | 0.81 | 0.74 | 1200 | 56 | EA y EM |
| <i>BA</i>   | 0.74 | 0.77 | 0.63 | 0.70 | 1200 | 56 | EA y EM |
| <i>K-VC</i> | 0.63 | 0.59 | 0.68 | 0.63 | 1200 | 56 | EA y EM |
| <i>NB</i>   | 0.93 | 0.96 | 0.90 | 0.93 | 1400 | 57 | EA y EM |
| <i>RL</i>   | 0.87 | 0.90 | 0.86 | 0.88 | 1400 | 57 | EA y EM |
| <i>MSV</i>  | 0.75 | 0.80 | 0.73 | 0.76 | 1400 | 57 | EA y EM |
| <i>BA</i>   | 0.68 | 0.80 | 0.56 | 0.66 | 1400 | 57 | EA y EM |
| <i>K-VC</i> | 0.64 | 0.66 | 0.69 | 0.67 | 1400 | 57 | EA y EM |
| <i>NB</i>   | 0.92 | 0.93 | 0.90 | 0.91 | 1600 | 58 | EA y EM |
| <i>RL</i>   | 0.86 | 0.85 | 0.86 | 0.85 | 1600 | 58 | EA y EM |
| <i>MSV</i>  | 0.76 | 0.72 | 0.81 | 0.76 | 1600 | 58 | EA y EM |
| <i>BA</i>   | 0.73 | 0.74 | 0.69 | 0.72 | 1600 | 58 | EA y EM |
| <i>K-VC</i> | 0.61 | 0.59 | 0.67 | 0.62 | 1600 | 58 | EA y EM |
| <i>NB</i>   | 0.92 | 0.93 | 0.92 | 0.92 | 1800 | 59 | EA y EM |
| <i>RL</i>   | 0.86 | 0.87 | 0.86 | 0.87 | 1800 | 59 | EA y EM |
| <i>MSV</i>  | 0.78 | 0.76 | 0.81 | 0.79 | 1800 | 59 | EA y EM |
| <i>BA</i>   | 0.75 | 0.77 | 0.72 | 0.75 | 1800 | 59 | EA y EM |
| <i>K-VC</i> | 0.59 | 0.57 | 0.73 | 0.64 | 1800 | 59 | EA y EM |
| <i>NB</i>   | 0.92 | 0.92 | 0.92 | 0.92 | 2000 | 60 | EA y EM |
| <i>RL</i>   | 0.89 | 0.88 | 0.89 | 0.88 | 2000 | 60 | EA y EM |
| <i>MSV</i>  | 0.81 | 0.79 | 0.83 | 0.81 | 2000 | 60 | EA y EM |
| <i>BA</i>   | 0.73 | 0.78 | 0.65 | 0.71 | 2000 | 60 | EA y EM |
| <i>K-VC</i> | 0.63 | 0.61 | 0.73 | 0.66 | 2000 | 60 | EA y EM |
| <i>RL</i>   | 0.90 | 0.88 | 0.88 | 0.88 | 200  | 61 | EM y EP |
| <i>NB</i>   | 0.88 | 0.88 | 0.82 | 0.85 | 200  | 61 | EM y EP |
| <i>BA</i>   | 0.70 | 0.65 | 0.65 | 0.65 | 200  | 61 | EM y EP |
| <i>K-VC</i> | 0.65 | 0.57 | 0.76 | 0.65 | 200  | 61 | EM y EP |
| <i>MSV</i>  | 0.63 | 0.53 | 1.00 | 0.69 | 200  | 61 | EM y EP |

|      |      |      |      |      |      |    |         |
|------|------|------|------|------|------|----|---------|
| NB   | 0.95 | 0.96 | 0.96 | 0.96 | 400  | 62 | EM y EP |
| RL   | 0.85 | 0.87 | 0.87 | 0.87 | 400  | 62 | EM y EP |
| MSV  | 0.85 | 0.88 | 0.84 | 0.86 | 400  | 62 | EM y EP |
| BA   | 0.83 | 0.97 | 0.71 | 0.82 | 400  | 62 | EM y EP |
| K-VC | 0.56 | 0.62 | 0.58 | 0.60 | 400  | 62 | EM y EP |
| NB   | 0.93 | 0.92 | 0.93 | 0.93 | 600  | 63 | EM y EP |
| RL   | 0.91 | 0.89 | 0.93 | 0.91 | 600  | 63 | EM y EP |
| MSV  | 0.83 | 0.84 | 0.80 | 0.82 | 600  | 63 | EM y EP |
| BA   | 0.81 | 0.81 | 0.82 | 0.81 | 600  | 63 | EM y EP |
| K-VC | 0.52 | 0.53 | 0.49 | 0.51 | 600  | 63 | EM y EP |
| NB   | 0.93 | 0.94 | 0.93 | 0.94 | 800  | 64 | EM y EP |
| RL   | 0.89 | 0.94 | 0.85 | 0.89 | 800  | 64 | EM y EP |
| MSV  | 0.82 | 0.91 | 0.75 | 0.82 | 800  | 64 | EM y EP |
| BA   | 0.76 | 0.89 | 0.65 | 0.75 | 800  | 64 | EM y EP |
| K-VC | 0.53 | 0.58 | 0.52 | 0.55 | 800  | 64 | EM y EP |
| NB   | 0.92 | 0.92 | 0.93 | 0.92 | 1000 | 65 | EM y EP |
| RL   | 0.89 | 0.86 | 0.92 | 0.89 | 1000 | 65 | EM y EP |
| MSV  | 0.81 | 0.78 | 0.88 | 0.83 | 1000 | 65 | EM y EP |
| BA   | 0.72 | 0.78 | 0.63 | 0.70 | 1000 | 65 | EM y EP |
| K-VC | 0.56 | 0.61 | 0.40 | 0.49 | 1000 | 65 | EM y EP |
| NB   | 0.91 | 0.93 | 0.89 | 0.91 | 1200 | 66 | EM y EP |
| RL   | 0.88 | 0.87 | 0.89 | 0.88 | 1200 | 66 | EM y EP |
| MSV  | 0.83 | 0.79 | 0.88 | 0.83 | 1200 | 66 | EM y EP |
| BA   | 0.70 | 0.73 | 0.60 | 0.66 | 1200 | 66 | EM y EP |
| K-VC | 0.63 | 0.61 | 0.60 | 0.60 | 1200 | 66 | EM y EP |
| NB   | 0.93 | 0.95 | 0.93 | 0.94 | 1400 | 67 | EM y EP |
| RL   | 0.90 | 0.91 | 0.91 | 0.91 | 1400 | 67 | EM y EP |
| MSV  | 0.86 | 0.90 | 0.84 | 0.87 | 1400 | 67 | EM y EP |
| BA   | 0.78 | 0.85 | 0.73 | 0.79 | 1400 | 67 | EM y EP |
| K-VC | 0.60 | 0.73 | 0.41 | 0.53 | 1400 | 67 | EM y EP |
| NB   | 0.92 | 0.92 | 0.91 | 0.92 | 1600 | 68 | EM y EP |
| RL   | 0.89 | 0.86 | 0.91 | 0.89 | 1600 | 68 | EM y EP |
| MSV  | 0.85 | 0.82 | 0.88 | 0.85 | 1600 | 68 | EM y EP |
| BA   | 0.77 | 0.80 | 0.69 | 0.74 | 1600 | 68 | EM y EP |
| K-VC | 0.65 | 0.68 | 0.53 | 0.59 | 1600 | 68 | EM y EP |
| NB   | 0.92 | 0.93 | 0.92 | 0.92 | 1800 | 69 | EM y EP |
| RL   | 0.88 | 0.86 | 0.91 | 0.89 | 1800 | 69 | EM y EP |
| MSV  | 0.85 | 0.83 | 0.89 | 0.86 | 1800 | 69 | EM y EP |
| BA   | 0.76 | 0.77 | 0.75 | 0.76 | 1800 | 69 | EM y EP |
| K-VC | 0.61 | 0.60 | 0.70 | 0.65 | 1800 | 69 | EM y EP |
| NB   | 0.92 | 0.91 | 0.92 | 0.92 | 2000 | 70 | EM y EP |
| RL   | 0.91 | 0.89 | 0.92 | 0.91 | 2000 | 70 | EM y EP |
| MSV  | 0.89 | 0.87 | 0.91 | 0.89 | 2000 | 70 | EM y EP |
| BA   | 0.75 | 0.75 | 0.75 | 0.75 | 2000 | 70 | EM y EP |
| K-VC | 0.64 | 0.61 | 0.74 | 0.67 | 2000 | 70 | EM y EP |

|      |      |      |      |      |      |    |             |
|------|------|------|------|------|------|----|-------------|
| RL   | 0.93 | 0.94 | 0.88 | 0.91 | 200  | 71 | EA, EM y EP |
| NB   | 0.88 | 0.88 | 0.82 | 0.85 | 200  | 71 | EA, EM y EP |
| BA   | 0.83 | 0.73 | 0.94 | 0.82 | 200  | 71 | EA, EM y EP |
| K-VC | 0.65 | 0.57 | 0.76 | 0.65 | 200  | 71 | EA, EM y EP |
| MSV  | 0.63 | 0.53 | 1.00 | 0.69 | 200  | 71 | EA, EM y EP |
| NB   | 0.95 | 0.96 | 0.96 | 0.96 | 400  | 72 | EA, EM y EP |
| RL   | 0.86 | 0.89 | 0.87 | 0.88 | 400  | 72 | EA, EM y EP |
| MSV  | 0.85 | 0.88 | 0.84 | 0.86 | 400  | 72 | EA, EM y EP |
| BA   | 0.79 | 0.85 | 0.76 | 0.80 | 400  | 72 | EA, EM y EP |
| K-VC | 0.56 | 0.62 | 0.58 | 0.60 | 400  | 72 | EA, EM y EP |
| NB   | 0.93 | 0.92 | 0.93 | 0.93 | 600  | 73 | EA, EM y EP |
| RL   | 0.91 | 0.89 | 0.93 | 0.91 | 600  | 73 | EA, EM y EP |
| MSV  | 0.83 | 0.84 | 0.80 | 0.82 | 600  | 73 | EA, EM y EP |
| BA   | 0.77 | 0.82 | 0.69 | 0.75 | 600  | 73 | EA, EM y EP |
| K-VC | 0.52 | 0.53 | 0.49 | 0.51 | 600  | 73 | EA, EM y EP |
| NB   | 0.93 | 0.94 | 0.93 | 0.94 | 800  | 74 | EA, EM y EP |
| RL   | 0.89 | 0.94 | 0.85 | 0.89 | 800  | 74 | EA, EM y EP |
| MSV  | 0.82 | 0.91 | 0.75 | 0.82 | 800  | 74 | EA, EM y EP |
| BA   | 0.76 | 0.84 | 0.70 | 0.76 | 800  | 74 | EA, EM y EP |
| K-VC | 0.53 | 0.58 | 0.52 | 0.55 | 800  | 74 | EA, EM y EP |
| NB   | 0.92 | 0.92 | 0.93 | 0.92 | 1000 | 75 | EA, EM y EP |
| RL   | 0.89 | 0.86 | 0.92 | 0.89 | 1000 | 75 | EA, EM y EP |
| MSV  | 0.81 | 0.78 | 0.88 | 0.83 | 1000 | 75 | EA, EM y EP |
| BA   | 0.76 | 0.80 | 0.72 | 0.76 | 1000 | 75 | EA, EM y EP |
| K-VC | 0.56 | 0.61 | 0.40 | 0.49 | 1000 | 75 | EA, EM y EP |
| NB   | 0.91 | 0.93 | 0.89 | 0.91 | 1200 | 76 | EA, EM y EP |
| RL   | 0.88 | 0.87 | 0.89 | 0.88 | 1200 | 76 | EA, EM y EP |
| MSV  | 0.83 | 0.79 | 0.88 | 0.83 | 1200 | 76 | EA, EM y EP |
| BA   | 0.75 | 0.77 | 0.67 | 0.71 | 1200 | 76 | EA, EM y EP |
| K-VC | 0.63 | 0.61 | 0.60 | 0.60 | 1200 | 76 | EA, EM y EP |
| NB   | 0.93 | 0.95 | 0.93 | 0.94 | 1400 | 77 | EA, EM y EP |
| RL   | 0.90 | 0.91 | 0.91 | 0.91 | 1400 | 77 | EA, EM y EP |
| MSV  | 0.86 | 0.90 | 0.84 | 0.87 | 1400 | 77 | EA, EM y EP |
| BA   | 0.77 | 0.86 | 0.69 | 0.76 | 1400 | 77 | EA, EM y EP |
| K-VC | 0.60 | 0.73 | 0.41 | 0.53 | 1400 | 77 | EA, EM y EP |
| NB   | 0.92 | 0.92 | 0.91 | 0.92 | 1600 | 78 | EA, EM y EP |
| RL   | 0.89 | 0.86 | 0.91 | 0.89 | 1600 | 78 | EA, EM y EP |
| MSV  | 0.85 | 0.82 | 0.88 | 0.85 | 1600 | 78 | EA, EM y EP |
| BA   | 0.76 | 0.80 | 0.66 | 0.72 | 1600 | 78 | EA, EM y EP |
| K-VC | 0.65 | 0.68 | 0.53 | 0.59 | 1600 | 78 | EA, EM y EP |
| NB   | 0.92 | 0.93 | 0.92 | 0.92 | 1800 | 79 | EA, EM y EP |
| RL   | 0.88 | 0.87 | 0.91 | 0.89 | 1800 | 79 | EA, EM y EP |
| MSV  | 0.85 | 0.83 | 0.89 | 0.86 | 1800 | 79 | EA, EM y EP |
| BA   | 0.74 | 0.79 | 0.66 | 0.72 | 1800 | 79 | EA, EM y EP |
| K-VC | 0.61 | 0.60 | 0.70 | 0.65 | 1800 | 79 | EA, EM y EP |

|      |      |      |      |      |      |    |             |
|------|------|------|------|------|------|----|-------------|
| NB   | 0.92 | 0.91 | 0.92 | 0.92 | 2000 | 80 | EA, EM y EP |
| RL   | 0.91 | 0.89 | 0.92 | 0.91 | 2000 | 80 | EA, EM y EP |
| MSV  | 0.89 | 0.87 | 0.91 | 0.89 | 2000 | 80 | EA, EM y EP |
| BA   | 0.76 | 0.81 | 0.68 | 0.74 | 2000 | 80 | EA, EM y EP |
| K-VC | 0.64 | 0.61 | 0.74 | 0.67 | 2000 | 80 | EA, EM y EP |
| NB   | 0.88 | 0.88 | 0.82 | 0.85 | 200  | 81 | Le          |
| RL   | 0.80 | 0.85 | 0.65 | 0.73 | 200  | 81 | Le          |
| BA   | 0.75 | 0.67 | 0.82 | 0.74 | 200  | 81 | Le          |
| K-VC | 0.60 | 0.53 | 0.47 | 0.50 | 200  | 81 | Le          |
| MSV  | 0.50 | 0.46 | 1.00 | 0.63 | 200  | 81 | Le          |
| NB   | 0.96 | 0.98 | 0.96 | 0.97 | 400  | 82 | Le          |
| RL   | 0.89 | 0.89 | 0.91 | 0.90 | 400  | 82 | Le          |
| BA   | 0.74 | 0.85 | 0.64 | 0.73 | 400  | 82 | Le          |
| MSV  | 0.69 | 0.88 | 0.51 | 0.65 | 400  | 82 | Le          |
| K-VC | 0.51 | 0.57 | 0.53 | 0.55 | 400  | 82 | Le          |
| NB   | 0.91 | 0.90 | 0.92 | 0.91 | 600  | 83 | Le          |
| RL   | 0.89 | 0.88 | 0.92 | 0.90 | 600  | 83 | Le          |
| BA   | 0.74 | 0.77 | 0.70 | 0.74 | 600  | 83 | Le          |
| MSV  | 0.71 | 0.75 | 0.64 | 0.69 | 600  | 83 | Le          |
| K-VC | 0.62 | 0.64 | 0.57 | 0.60 | 600  | 83 | Le          |
| NB   | 0.91 | 0.94 | 0.90 | 0.92 | 800  | 84 | Le          |
| RL   | 0.85 | 0.88 | 0.84 | 0.86 | 800  | 84 | Le          |
| BA   | 0.73 | 0.87 | 0.60 | 0.71 | 800  | 84 | Le          |
| MSV  | 0.70 | 0.82 | 0.60 | 0.69 | 800  | 84 | Le          |
| K-VC | 0.59 | 0.64 | 0.60 | 0.62 | 800  | 84 | Le          |
| NB   | 0.89 | 0.89 | 0.90 | 0.90 | 1000 | 85 | Le          |
| RL   | 0.84 | 0.84 | 0.85 | 0.84 | 1000 | 85 | Le          |
| MSV  | 0.76 | 0.76 | 0.78 | 0.77 | 1000 | 85 | Le          |
| BA   | 0.73 | 0.80 | 0.63 | 0.71 | 1000 | 85 | Le          |
| K-VC | 0.56 | 0.57 | 0.58 | 0.57 | 1000 | 85 | Le          |
| NB   | 0.90 | 0.92 | 0.86 | 0.89 | 1200 | 86 | Le          |
| RL   | 0.83 | 0.83 | 0.82 | 0.82 | 1200 | 86 | Le          |
| MSV  | 0.74 | 0.69 | 0.81 | 0.74 | 1200 | 86 | Le          |
| BA   | 0.70 | 0.72 | 0.61 | 0.66 | 1200 | 86 | Le          |
| K-VC | 0.63 | 0.59 | 0.68 | 0.63 | 1200 | 86 | Le          |
| NB   | 0.93 | 0.96 | 0.90 | 0.93 | 1400 | 87 | Le          |
| RL   | 0.87 | 0.90 | 0.86 | 0.88 | 1400 | 87 | Le          |
| MSV  | 0.75 | 0.80 | 0.73 | 0.76 | 1400 | 87 | Le          |
| BA   | 0.69 | 0.80 | 0.56 | 0.66 | 1400 | 87 | Le          |
| K-VC | 0.64 | 0.66 | 0.69 | 0.67 | 1400 | 87 | Le          |
| NB   | 0.92 | 0.93 | 0.90 | 0.91 | 1600 | 88 | Le          |
| RL   | 0.86 | 0.85 | 0.86 | 0.85 | 1600 | 88 | Le          |
| MSV  | 0.76 | 0.72 | 0.81 | 0.76 | 1600 | 88 | Le          |
| BA   | 0.71 | 0.74 | 0.61 | 0.67 | 1600 | 88 | Le          |
| K-VC | 0.61 | 0.59 | 0.67 | 0.62 | 1600 | 88 | Le          |

|             |      |      |      |      |      |    |    |
|-------------|------|------|------|------|------|----|----|
| <i>NB</i>   | 0.92 | 0.93 | 0.92 | 0.92 | 1800 | 89 | Le |
| <i>RL</i>   | 0.86 | 0.87 | 0.86 | 0.87 | 1800 | 89 | Le |
| <i>MSV</i>  | 0.78 | 0.76 | 0.81 | 0.79 | 1800 | 89 | Le |
| <i>BA</i>   | 0.74 | 0.77 | 0.69 | 0.73 | 1800 | 89 | Le |
| <i>K-VC</i> | 0.59 | 0.57 | 0.73 | 0.64 | 1800 | 89 | Le |
| <i>NB</i>   | 0.92 | 0.92 | 0.92 | 0.92 | 2000 | 90 | Le |
| <i>RL</i>   | 0.89 | 0.88 | 0.89 | 0.88 | 2000 | 90 | Le |
| <i>MSV</i>  | 0.81 | 0.79 | 0.83 | 0.81 | 2000 | 90 | Le |
| <i>BA</i>   | 0.73 | 0.76 | 0.65 | 0.70 | 2000 | 90 | Le |
| <i>K-VC</i> | 0.63 | 0.61 | 0.73 | 0.66 | 2000 | 90 | Le |