

Vea discusiones, estadísticas y perfiles de autores para esta publicación en: <https://www.researchgate.net/publication/351285564>

Detección de noticias falsas basada en la incrustación de documentos y palabras usando una máquina clasificadores de aprendizaje

Artículo en Revista de tecnología de la información teórica y aplicada · Abril de 2021

CITACIONES

0

LEE

72

4 autores, incluyendo:



IbrahimEldesoky

Universidad Beni Suef

10 PUBLICACIONES 6 CITACIONES

VER EL PERFIL

Algunos de los autores de esta publicación también están trabajando en estos proyectos relacionados:



Project

Herramienta de anotación de texto médico basada en IBM Watson Platform Ver Proyecto

DETECCIÓN DE FALSAS NOTICIAS EN BASE A LA INTEGRACIÓN DE PALABRAS Y DOCUMENTOS UTILIZANDO EL APRENDIZAJE MÁQUINA CLASIFICADORES

¹IBRAHIM EL DESOUKY FATTOH * ²FARID ALI MOUSA

¹ Departamento de Ciencias de la Computación, Facultad de Computación e Inteligencia Artificial, Universidad Beni-Suef, Egipto.

² Departamento de Tecnología de la Información, Facultad de Computación e Inteligencia Artificial, Beni-Suef Universidad, Egipto.

Correo electrónico: ¹Ibrahim_desoky@fcis.bsu.edu.eg, ²faly@msa.eun.eg

ABSTRACTO

Las noticias falsas son un problema que tiene un efecto importante en nuestra vida. La detección de fake news se considera un área de investigación interesante que presenta alguna limitación de los recursos disponibles. En esta investigación, proponemos un modelo de clasificación que es capaz de detectar noticias falsas basado en la incrustación de Doc2vec y Word2vec como métodos de extracción de características. En primer lugar, comparamos entre los dos enfoques utilizando diferentes algoritmos de clasificación. Según los experimentos aplicados, la clasificación basada en el modelo Doc2vec arrojó resultados prometedores con más de un clasificador. La máquina de vectores de soporte resultó la mejor precisión con 95,5% **seguido de Regresión logística 94,7%** y la memoria a largo y corto plazo produjo la precisión más baja. Por otro lado, el modelo de incrustación de Word2vec basado en la clasificación da como resultado una alta precisión solo con el clasificador Long Short TermMemory con un 94,3%. En segundo lugar, los modelos de clasificación basados en el Doc2vec propuesto han demostrado que superan a un modelo correspondiente que se basa en TF-IDF en el mismo conjunto de datos utilizando clasificadores de Regresión Logística y Máquina de Vector de Soporte.

Palabras clave: *Detección de noticias falsas; Word2Vec; Doc2Vec; Aprendizaje automático; Aprendizaje profundo*

1. INTRODUCCIÓN

Con el lanzamiento de la World Wide Web y alta adquisición de redes sociales (por ejemplo, Facebook y Twitter), el intercambio de información ha abierto un camino nunca antes visto en la historia de la humanidad.

Además de otras aplicaciones, los quioscos se han beneficiado del uso extensivo de los sitios de redes sociales al proporcionar a sus suscriptores noticias actualizadas casi en tiempo real. Los medios se expandieron desde periódicos, revistas y tabloides a medios modernos como plataformas de noticias de Internet, sitios web, transmisiones de redes sociales, etc. Esto, lamentablemente, se está utilizando para engañar a los lectores y difundir fácilmente noticias falsas que normalmente contienen información tergiversada o incluso **fotografías fabricadas**. [1]

Difundir noticias falsas puede tener importantes consecuencias negativas e incluso importantes asuntos públicos. puede verse influenciado o incluso distorsionado. En los últimos años, las noticias falsas, especialmente después de las elecciones estadounidenses de 2016, han ganado más atención. Las noticias falsas son difíciles de detectar para las personas. La única forma en que una persona puede detectar noticias falsas manualmente es obteniendo una inmensa comprensión del tema que se está cubriendo. Y si

la información del informe es verdadera o falsa, es muy difícil de identificar de manera efectiva. [2]

Las noticias falsas a menudo se describen como hechos precisos, pero en realidad no existe. Asumimos en el mundo de hoy que lo que leemos en blogs o redes sociales es cierto y no buscamos verificar si la información brindada es precisa o no. Dado que la cultura, los individuos, las tecnologías y los sistemas modernos hacen todos estos intentos, todavía vemos todos los días, de alguna forma o forma, algunas noticias falsas. [3]

La detección de noticias falsas puede ser muy útil para mejorar el medio ambiente en muchos aspectos en un escenario de enormes cantidades de conocimiento de noticias falsas. Cuando profundizamos en este tema de las noticias falsas, hay dos amplios campos de inspiración, uno es la identificación de las noticias falsas. noticias y el segundo es la clasificación de las noticias falsas y su amplia variedad. [4]

El objetivo clave es identificar las noticias falsas, que es una cuestión de clasificación de texto convencional con una propuesta inmediata. El modelo es importante para discernir entre noticias "reales" y "falsas". Esto se traduce en implicaciones para las plataformas de redes sociales como Facebook, Instagram, Twitter y para las aplicaciones de mensajería instantánea como WhatsApp, donde esta

Las noticias reciben un gran impulso y se vuelven virales en todo el país y en todo el mundo. El método sugerido ayuda a encontrar la credibilidad de las noticias. Si la noticia no es precisa, el informe de noticias recomendaría al lector con el artículo de noticias relevante [5].

En este artículo, proponemos usar tanto la incrustación de palabras (Word2vec) como la incrustación de documentos (Doc2vec) como métodos de extracción de características para la creación de varias clasificaciones, como regresión logística, vector de soporte máquinas, bosques aleatorios, redes neuronales con perceptrón multicapa y clasificadores de memoria a corto plazo. Los clasificadores sugeridos utilizados para diferenciar entre hechos engañosos y noticias reales. Los clasificadores se evalúan y educan en un conjunto de datos basado en la web.

Este documento se prepara de la siguiente manera: sección 2 cubrirá el área de trabajo relacionado, la sección 3 discutirá las fases del modelo propuesto, la sección 4 cubrirá el conjunto de datos utilizado, experimentos aplicados y analiza los resultados obtenidos, y la sección 5 proporciona la conclusión y el trabajo futuro.

2. TRABAJO RELACIONADO

Mumbai publicó tres artículos de investigación sobre Identificación de noticias falsas por parte de tres estudiantes del Instituto de Tecnología de la Sociedad de Educación de Vivekananda en 2018. Escribieron que la era de las redes sociales comenzó en el siglo XX en su artículo de investigación. Con el tiempo, el uso de la web se expande, aumentan las publicaciones, aumenta el número de artículos. Utilizaron diversos métodos y herramientas para identificar noticias falsas, como técnicas de procesamiento del lenguaje natural (PNL), inteligencia artificial (IA) y aprendizaje automático. (ML). [6]. Nguyen Vo, de la Universidad Tecnológica de la ciudad de Ho Chi Minh (HCMUT) en Camboya, investigó y aplicó la identificación de noticias falsas en 2017. En su proyecto, identificación de noticias falsas, utilizó GRU bidireccional. Yang y col. utilizó algunos algoritmos de aprendizaje profundo además de otros modelos de aprendizaje profundo, como Auto-Encoders, CNN y GAN. [7]

Rubin y col. [8] recomendó un modelo para la detección de noticias sobre sátira y comedia. Revisaron e inspeccionaron 360 artículos de noticias en satírico, principalmente en cuatro áreas: cívica, científica, del mercado y de las noticias blandas. Sugirieron un modelo de clasificación de máquinas vectoriales de soporte con cinco características clave que se construyeron sobre la base de su estudio satírico de noticias. Hay cinco funciones; Absurdo, Gramática, Humor, Puntuación y Afecto Negativo. Su precisión general del 90 por ciento se ha alcanzado con solo tres combinaciones de atributos; Absurdo,

Puntuación y gramática. Los investigadores de [9] exploraron los usos del término, especialmente con respecto a las noticias falsas. Se han encontrado seis definiciones amplias de la palabra "noticias falsas", tales como; sátira de noticias, fabricación, parodia de noticias, manipulación (por ejemplo, fotografía), publicidad (por ejemplo publicidad presentada como periodismo) y propaganda. Sus objetivos y asignaciones eran "la apariencia de una noticia real". Fallis [10] explora cómo se ha descrito la desinformación. Concluyó que "la desinformación es información engañosa con una función engañosa". En [8], Rubin divide las noticias falsas en tres categorías separadas: fabricación extrema, enorme falsificaciones y falsedades humorísticas. En lugar de cualquier otra agrupación, no explican por qué eligen estos tipos. Sin embargo, discutieron en **detalla lo que contendrá cada grupo y cómo discriminar entre ellos**. Destacan además la ausencia de un organismo para realizar tal estudio y subrayan los 9 criterios sobre la creación de tal estudio. entidad que son; "Entrada en formato de texto digital, 'Hechos básicos verificables', 'Marco de tiempo predefinido', 'Lenguaje y cultura', 'La forma de entregar noticias', 'Preocupaciones pragmáticas' y 'Homogeneidad de las longitudes y la redacción son importantes. Probando con el desafío de clasificación, Kai Shu, Amy Sliva, Suhang Wang y Jiliang Tang [11] sugirieron características como; una cantidad de caracteres por palabra, cantidad de palabras, tasas de frase y oraciones, secciones de etiquetas de habla (es decir, n-gramas y enfoques de bolsa de palabras).

En el artículo de investigación, los autores de [12] propusieron la detección del fraude con un conjunto de datos de referencia denominado "LIAR", y apareció un aumento aparente en la eficiencia de la identificación de mensajes / noticias falsos. Los autores argumentaron que el corpus se utilizó para identificar lugares, vistas de minería, identificación de rumores y análisis sobre políticas de PNL.

Los investigadores de [13] implementaron la necesidad de detectar engaños. Utilizaron el enfoque de ML integrando la publicidad con enfoques de contenido social. Los investigadores afirmaron que el estudio funcionó bien en comparación con la literatura mencionada. Usaron el chatbot de Facebook Messenger para ejecutarlo. Se incluyen tres numerosos conjuntos de datos de publicaciones de noticias italianas de Facebook. Los algoritmos booleanos de crowdsourcing son todos enfoques orientados al contenido que se implementaron con señales sociales y de contenido. Los autores identificaron la tabloidización en forma de cebo de clics en [14]. El cebo de clics se describió como un método de propagación rápida en línea de rumores e información errónea. Como medio de insatisfacción, los académicos han explorado posibles métodos para detectar automáticamente los clickbait.

Las métricas de contenido, incluido el estándar teórico léxico y semántico, fueron implementadas por los autores. En [15], los autores han estudiado y analizado el alcance y el desempeño respectivos, y se utilizaron valores, herramientas y algoritmos para distinguir las noticias falsificadas de las fabricadas. El documento también describió las dificultades del estudio por las características desconocidas de las noticias falsas y los complejos vínculos entre las noticias, los escritores y los temas.

Los autores discutieron el paradigma del detector de falsos para la inferencia automática de noticias falsas. Depende de la identificación textual y proporciona un paradigma de red de amplio alcance para estudiar simultáneamente las representaciones de noticias, escritores y temas [16]. Fake-Detector aborda dos componentes clave: la representación de aprendizaje de características y la inferencia de marcas de reputación, que juntas forman el modelo de red profundamente difundido de Fake-Detector. Otra investigación de William Yang et.al. en [17], debido a la falta de datos falsos y su eficacia. Decidió introducir un nuevo conjunto de datos 'LIAR'. Es un conjunto de información de identificación de noticias de libre acceso para las noticias falsas. Utilizándolo utilizando una amplia variedad de métodos como regresión logística, admite modelos de máquina vectorial y aprendizaje profundo (Bidirectional-LSTM y Convolutional Neural Networks). Los hallazgos mostraron que los modelos de CNN son de hecho los mejores. Himank Gupta. et. a l. [18] tiene un sistema enfocado en un enfoque particular de aprendizaje automático que resuelve una serie de problemas, incluida la escasez de precisión, el retraso (Bot-Maker) y el cálculo rápido del tiempo para miles de tweets en un segundo. Primero, se obtuvieron 400,000 tweets de conjuntos de datos HSpam14. Luego, se caracterizan aún más los 150.000 mensajes no deseados y los 250.000 tuits que no son spam. También extraen algunas características más ligeras junto con las 30 palabras principales que le dan al modelo Bag-of-Words la mayor ganancia de información. Alcanzaron el 91,65% lapso de tiempo (Bot-Maker) y cálculo de tiempo rápido para miles de tweets en un segundo. Primero, se obtuvieron 400,000 tweets de conjuntos de datos HSpam14. Luego, se caracterizan aún más los 150.000 mensajes no deseados y los 250.000 tuits que no son spam. También extraen algunas características más ligeras junto con las 30 palabras principales que le dan al modelo Bag-of-Words la mayor ganancia de información. Alcanzaron el 91,65% lapso de tiempo (Bot-Maker) y cálculo de tiempo rápido para miles de tweets en un segundo. Primero, se obtuvieron 400,000 tweets de conjuntos de datos HSpam14. Luego, se caracterizan aún más los 150.000 mensajes no deseados y los 250.000 tuits que no son spam. También extraen algunas características más ligeras junto con las 30 palabras principales que le dan al modelo Bag-of-Words la mayor ganancia de información. Alcanzaron el 91,65%

precisión y superó el enfoque actual en casi un 18%. El autor se refirió a Kai Shu et al. en [11], para una encuesta más detallada del trabajo de identificación de noticias falsas tratando de concentrarse en un sistema de identificación de cebos de clic ligero centrado en la funcionalidad de título de alto nivel.

3. MODELO PROPUESTO

3.1 Preprocesamiento de datos

Hasta crear un modelo vectorial, los datos deben estar sujeto a algunas mejoras, como la eliminación de palabras vacías, tokenización, segmentación en minúsculas y eliminación de puntuaciones. Esto ayuda a reducir el tamaño de los datos reales al eliminar los datos innecesarios. Para cada documento, construimos un

característica estandarizada de manejo para eliminar signos de puntuación y características que no sean letras.

Detener la eliminación de palabras: Las palabras vacías en un idioma son palabras sin sentido que producen ruido cuando se usan como características de clasificación de texto.

Hay palabras que se encuentran ampliamente en frases para vincular mejor los pensamientos o la estructura de la expresión. Las palabras vacías se conocen como publicaciones, preparaciones, conjunciones y pronombres. Los términos populares se han eliminado de cada documento, como; sobre, a, an, como, están en, por, ser, desde, para, en, es, en, cómo, o, eso, esto, estos, fue, cuándo, dónde, quién, qué, también, etc. Después de eso, los documentos procesados se conservaron y se transmitieron a la siguiente etapa.

Derivado: El siguiente paso es cambiar los tokens después de la tokenización a una forma regular de datos. La derivación esencialmente traduce los términos a la forma original y reduce el número de tipos de palabras o clases dentro de los resultados. Los términos "correr", "corrió" o "corredor" se reducirán a "correr". Nuestro rastreo se utiliza para una clasificación más fácil y eficaz. Además, utilizamos el lematizador Porter debido a su precisión, ya que es el algoritmo de lematización más utilizado.

3.2 Extracción de características

El aprendizaje de alta dimensión es uno de los desafíos de la categorización. descripción del texto. Los documentos contienen un gran número de expresiones, frases y palabras que hacen que el proceso de aprendizaje se vuelva altamente computacional. Además, la precisión y la salida del clasificador pueden verse influenciadas por funciones irrelevantes y redundantes. Es mejor reducir el tamaño de la función y evitar mediciones de área amplia de la función. Las integraciones para la mayoría de nuestros modelos se generan con los modelos Word2Vec y Doc2Vec.

3.2.1 Incrustación de Word2vec:

Word2vec fue desarrollado para "Representación de la computación vectorial continua de palabras a partir de conjuntos de datos más grandes" por Tomas Mikolov y el equipo de Google en 2013 [19]. Querían producir un aprendizaje más rápido modelos, que eran comunes en ese momento y claramente más fácil, de una manera diferente a los modelos de redes neuronales.

El autor y su equipo han propuesto dos modelos predictivos: el CBOW (Continuous Bag of Words) y el Skip-gram. Los modelos predictivos entrenan sus vectores para potenciar sus capacidades predictivas, de modo que se aprendan mejores resultados.

Bolsa de palabras continua (CBOW):

Las palabras en este modelo se toman, por ejemplo, luego se suman todos los vectores y la relación entre las palabras en una oración para predecir la palabra que falta en una oración aplicando la misma proyección.

ventana. La razón por la que lo llaman bolsa de palabras es que no es necesario ordenar las palabras.

Saltar gramo: El modelo inverso del modelo CBOW. La capa de entrada contendrá la palabra de destino y en la capa de salida, el contexto de la palabra. Pero, ¿cómo adivinas la frase a partir de una palabra? Primero, se necesitarán palabras de vocabulario de los documentos de capacitación. En segundo lugar, insertando la palabra que queremos usar en el vector one-hot, luego colocaremos 1 para la palabra usada y 0 para el resto. Después de eso, multiplicamos el vector one-hot con una matriz de frases potenciales y nuestro vector one-hot se ajustará al correcto.

3.2.2 Incorporación de Doc2vec:

Doc2vec es una técnica que utiliza un vector para describir un texto que es una versión simplificada de Word2vec incrustado en el procesamiento del lenguaje natural. Los modelos de Bolsa de palabras distribuida (DBOW) y Memoria distribuida (DM) Doc2vec sugirió. DBOW es un modelo analógico de doc2vec para word2vec skip-gram. En el sentido actual o en el ensayo, la mente DM recuerda lo que se ha perdido. reducirá los problemas de BOW al mínimo [20]. Doc2VecC consta de una capa de entrada, una capa de pantalla y una capa de salida que pronostica la palabra de destino. Idéntico a Word2Vec. Las palabras incrustadas de los documentos adyacentes tienen un significado local, mientras que el vector interpreta todo el documento como un contexto global. A diferencia de los vectores de párrafo, que aprenden explícitamente un vector único para cada artículo, Doc2VecC es el promedio de las incrustaciones de palabras muestreadas aleatoriamente en el texto. Cada párrafo se asigna en un DM a un solo vector descrito por una columna de matriz D y cada término está vinculado a un vector especial, que se expresa mediante una columna de matriz W. Los vectores de párrafo y frase se promedian o concatenan para pronosticar en contexto la siguiente palabra. En los experimentos se utiliza la concatenación como método para integrar vectores. Puede usarse como otro término en el símbolo de párrafo. Sirve como un recuerdo que recuerda lo que se perdió, o el tema del párrafo.

Otro enfoque es el modelo DBOW para ignorar los términos contextuales de entrada pero empujar el modelo para predecir palabras muestreadas al azar en la salida del párrafo. En particular, en cada iteración del descenso del gradiente estocástico, se toma una ventana de texto, se toma una palabra aleatoria de la ventana de texto y una Se crea la tarea de clasificación, siempre que el vector en párrafos.

3.3 Detección de noticias falsas

Los pasos básicos del modelo propuesto para detección de noticias falsas y clasificadores utilizados en esta investigación que se muestran en la figura 1. El primer paso es preprocesar el texto y luego crear los vectores de características

utilizando dos métodos diferentes (Doc2vec y Word2vc). Usamos para detectar la clase del nuevo documento cinco clasificadores. Los clasificadores son, Máquina de vectores de soporte (SVM), Red neuronal artificial (ANN), Memoria a corto plazo a largo plazo (LSTM), Regresión logística (LR), Perceptrón multicapa (MLP) y Bosque aleatorio (RF). Word2vec y el modelo DM de Doc2vec se utilizan como paso de extracción de funciones para SVM, LSTM, MLP, LR y RF; las funciones extraídas se utilizan para entrenar

los clasificadores y luego probar cada modelo contra el conjunto de prueba.

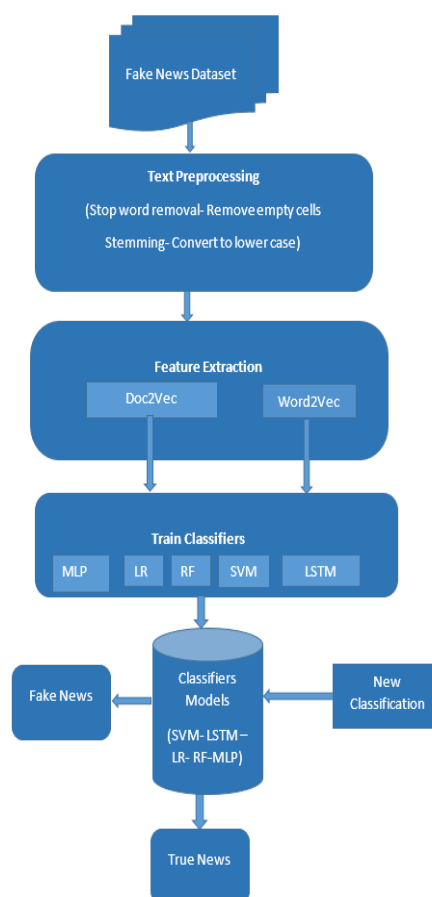


Figura 1. Modelo propuesto de detección de noticias falsas

La siguiente es una breve descripción de la clasificadores aplicados

3.3.1 Regresión logística

Logístico la regresión es tal a bien-método de clasificación conocido. Las variables de función se sabe que no son aleatorios en la edición más simple. La respuesta de clase es una variable aleatoria binaria que toma un valor de 1 para la clase de interés con

una cierta probabilidad p y un valor de 0 con probabilidad $1 - p$. La probabilidad de éxito " p " es una función de valores de las variables de características, de hecho, los registros impares "Iniciar sesión" o la razón logaritmo de probabilidades del

las variables predictoras es una función lineal.

La regresión logística incluye la verificación de hipótesis, junto con otras evaluaciones, cálculos y medidas de aptitud sobre el valor de cada variable. La verificación de importancia variable se puede utilizar para seleccionar características en la configuración de clasificación. Las implementaciones informáticas modernas implican varias variaciones de selección de variables escalonadas (iterativas).

La clasificación logística podría ser el método de minería de datos más comúnmente utilizado debido a la analogía matemática con la regresión múltiple ordinaria y la simplicidad de la recopilación automática de variables. [21]

Componentes de a probabilístico máquina algoritmo de aprendizaje: la regresión logística, como el ingenuo Bayes, es una prueba probabilística graduada que utiliza instrucción supervisada. Se requiere un grupo de entrenamiento de pares de entrada / salida para clasificadores de máquina ($x_{(i)}$; $y_{(i)}$). Podemos usar superíndices entre paréntesis para cada instancia en la colección de capacitación; cualquier caso podría ser un documento de clasificación individual:

1. La representación de la entidad de entrada. Este es un vector de características para cualquier entrada de observación $x_{(i)}$, [$X_1; X_2; \dots; X_{\text{norte}}$]. Vamos a enumerar la característica de entrada i , $x_{(i)}$ como $x_{(i)}$ a menudo simplificado como x_i , sin embargo, notaremos la notación f_i , $f_i(x)$, o para la clasificación de la multiclase, $f_i(c; x)$.
2. Una función de clasificación determinada por $p(y | x)$ para Y para la clase aproximada. Las herramientas de clasificación Sigmoid y SoftMax se agregan en el siguiente sección.
3. Función de aprendizaje objetivo que normalmente elimina errores en escenarios de entrenamiento.
4. Un algoritmo para la optimización de la función objetivo (ver Función sigmoidea en la Figura 2). Implementamos el algoritmo de gradiente descendente estocástico. La regresión logística tiene dos etapas: la primera es la fase de entrenamiento: el método se entrena con gradiente estocástico hacia abajo y pérdida de entropía cruzada (específicamente ponderaciones w y b). Proporcionando el ejemplo de prueba x , se determina $p(y|x)$ y se devuelve la marca $y = 1$ o $y = 0$.

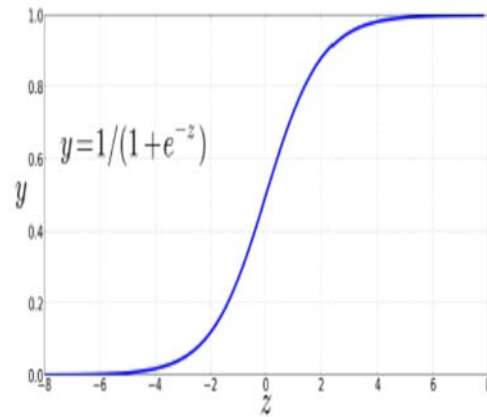


Figura 2. La función sigmoidea de la regresión lineal

3.3.2 Clasificador de bosque aleatorio

Un bosque aleatorio es un metaestimador compuesto por varios árboles de decisión que funcionan como un conjunto. Al construir cada árbol individual, utiliza el bagaje y la aleatoriedad para intentar crear un bosque de árboles correlacionados negativamente cuya predicción del comité es más confiable que cualquier árbol individual. Cada Un solo árbol difunde una prevención de clase al azar. bosque, y la clase más votada se convierte en el predictor del modelo.

El modelo de bosque aleatorio funciona bien, porque un gran número de árboles (modelo) correlacionados comparativamente negativamente que funcionan como un comité supera a los modelos constituyentes individuales, así como porque la baja asociación entre modelos es el secreto de este modelo forestal aleatorio. [21] Los bosques aleatorios son un método completo importante basado en el algoritmo CART para el cultivo de árboles individuales. La idea principal es incorporar varios árboles en un conjunto robusto en lugar de solo construir un libro de decisiones para la predicción. Los bosques aleatorios utilizan ensacado para plantar numerosos árboles tomando una cantidad significativa de muestras de datos de entrenamiento, es decir, mediante muestreo y sustitución, muestras que son del mismo tamaño que los datos de entrenamiento. Los datos muestreados se envían luego a un algoritmo similar a CART para construir para cada muestra de arranque, un árbol de decisión. Dado que las muestras de este bootstrap comprenden varias secciones de los datos originales, los árboles variarán entre muestras y formarán una colección de árboles separados.

3.3.3. Perceptrón multicapa (MLP)

Un perceptrón multicapa (MLP) es un modelo de red neuronal artificial de retroalimentación, donde los conjuntos de datos de entrada se asignan a una colección de salidas adecuadas. Comprende formas de tres capas; la entrada, la salida y la capa oculta. La capa de entrada recibe la señal de procesamiento.

La tarea adecuada se realiza a través de la capa de salida, incluida la predicción y la clasificación. El verdadero motor de procesamiento del MLP consiste en un número infinito de capas ocultas entre el lado de entrada y el de salida. Como una red de transmisión en un MLP, los datos fluyen en una dirección hacia adelante desde la capa de entrada a la capa de salida, como la red neuronal de alimentación hacia adelante.

Las neuronas MLP están condicionadas por la red neuronal de retropropagación. Los MLP se han desarrollado para aproximar cualquier función continua y para resolver problemas que no pueden aislarse linealmente.

Descripción del patrón, identificación, la estimación y la aproximación son los casos clave de la aplicación MLP [22].

El clasificador ANN consta de nodos, una capa de entrada como (x_1, x_2, \dots, x_n) y una capa oculta opcional, así como una capa de salida (y) , (que se muestra en la Figura 3). El objetivo de la ANN es analizar un conjunto de ponderaciones (w) (entre los nodos de entrada, nodos ocultos y nodos de salida) que minimizan el error de la suma cuadrada total.

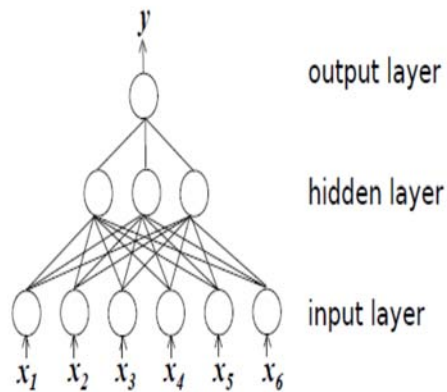


Figura 3. Red neuronal artificial

Los pesos w están sintonizados con respecto al parámetro de aprendizaje $\lambda \in [0, 1]$ durante el entrenamiento, hasta que todos los resultados sean consistentes con el resultado. Se pueden hacer cambios de peso demasiado drásticos si λ es alto, mientras que pueden ser necesarias más iteraciones (llamadas épocas) si el valor es demasiado pequeño antes de que el modelo aprenda lo suficiente de los datos de entrenamiento. La especificación de parámetros que se aprenden de

Se sabe que los datos de entrenamiento (es decir, el stock de datos de entrenamiento) son difíciles de usar ANN y, por lo tanto, funcionan mal con datos invisibles. Si hay demasiados nodos ocultos, el sistema podría sobrescribir los datos actuales y evitar que el sistema actualice con precisión el valor de entrada si hay muy pocos. Además, es importante elegir un criterio de parada. Esto puede implicar detenerse cuando la pérdida acumulada de la red está por debajo de los valores predeterminados o si se realizan varias épocas (iteraciones).

3.3.4 Memoria a corto plazo (LSTM)

La memoria a largo plazo a corto plazo (LSTM) [16] es un tipo especializado de redes neuronales recurrentes que tienen la capacidad de registrar patrones de larga duración de forma selectiva. Es una excelente alternativa para modelar datos secuenciales y, por tanto, para estudiar dinámicas complejas del comportamiento humano. El estado celular se considera memoria a largo plazo. Dado que las celdas son de naturaleza recursiva, los datos anteriores se pueden almacenar en ellas. Para cambiar la condición de la celda, se usa la puerta de olvido debajo del estado de la celda. Los valores de liberación de la puerta olvidada determinan qué datos olvidar multiplicando cero a una posición en la matriz. El conocimiento se conserva dentro de la celda, mientras que la salida de la puerta de olvido es una. La entrada de puertas específicas la información, los estados de la celda deben ingresar. Finalmente, en la puerta de salida, indique que la información debe ser cambiado al estado secreto oculto.

LSTM se usa comúnmente en aplicaciones de múltiples modelos, incluido el subtítulo de fotos, y también mejora las redes neuronales recurrentes (RNN) de vainilla. Está diseñado para garantizar una dependencia duradera y reducir sustancialmente el problema de la desaparición progresiva de las enfermeras registradas. LSTM lee estas palabras uno a uno y retiene un estado de memoria m_t en las dimensiones D y un estado oculto h_t en D , dada una serie de palabras $\{x_1, x_2, \dots, x_n\}$. [23] La figura 4 muestra el diagrama LSTM

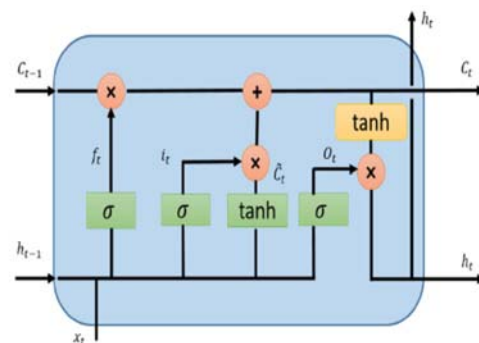


Figura 4. Diagrama de memoria a corto plazo (LSTM)

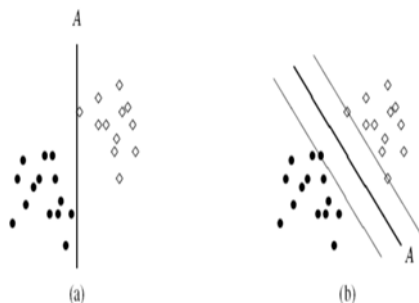
3.3.5 Máquina de vectores de soporte (SVM)

Debido a las capacidades de las máquinas vectoriales de soporte, los datos de alta dimensión se manejan de manera efectiva. SVM es una descripción de la discriminación descrita por una separación de hiperplano. Y si SVM en sí es un clasificador de dos, también se puede utilizar para la clasificación múltiple.

Opera para muchas clases aplicando una técnica versus una para cada par de clases. El propósito clave del algoritmo es identificar instancias que dependen de una ecuación lineal. También puede realizar una clasificación no lineal con una función de kernel. La clasificación se suministra con instancias premarcadas y la SVM busca un hiperplano que maximice el margen eligiendo puntos como soporte vectores.

Support Vector Machines (SVM), que son un básico concepto pero efectivo, son una de las técnicas más comunes de discriminación lineal. Lo siguiente funciona SVM: mapear muestras desde el campo de entrada a un espacio de función de alta dimensión con el fin de localizar el hiperplano 'correcto' mientras elige el muestras. Si su margen es mayor, un hiperplano que divide H, es mayor. El margen de los dos hiperplanos en ambos lados paralelos a H es la media distancia sin tramos de muestra. De la teoría de la evitación del riesgo (una estimación de la función de falla predicha, es decir, las muestras de la clasificación errónea), cuanto mayor es el margen, mejor es el error de generalización del clasificador.

¡Error! No se encontró la fuente de referencia. ilustrar esta noción. Es evidente que el mismo conjunto de entrenamiento tiene un hiperplano de división diferente. El hiperplano que divide las muestras de las clases más cercanas a mayor distancia de sus elegidas, siempre que los dos grupos de muestras estén separados por el mayor margen entre sí y, por lo tanto, sean menos propensos a fallas pequeñas en la dirección de la



hiperplano. [24, 25]

Figura 5. Máquina de vectores de soporte

4. CONJUNTO DE DATOS, EXPERIMENTOS, Y RESULTADOS

4.1 Conjunto de datos

Dos artículos diferentes contienen noticias falsas y reales del conjunto de datos utilizado en este estudio. La recopilación del conjunto de datos se realizó a partir de fuentes del mundo real; cierto Se recopilaron artículos del sitio web de noticias Reuters.com. Si bien los artículos de las noticias falsas se han compilado de numerosos medios, como sitios web poco confiables marcados por PolitiFact (una organización de verificación de datos de EE. UU.) Y Wikipedia. El conjunto de datos incluye varios tipos de documentos. Sin embargo, la mayoría se centra en temas de actualidad política y mundial.

En [26, 27] se describe más información sobre el conjunto de datos.

4.2 Métricas de evaluación

La Tabla 1 muestra las cuatro evaluaciones básicas métricas descripción utilizada para evaluar los resultados de la experimentos aplicados

Tabla 1 Misericordias de la evaluación

	Fórmula	Evaluación Atención
Precisión (acc)	$\frac{tp}{tp + fp}$	Típicamente calcula los porcentaje de pronósticos precisos sobre el total número de Medido instancias
Sensibilidad (sn)	$\frac{tp}{tp + fn}$	Calcular precisamente fracciones clasificadas de patrones positivos
Especificidad (sp)	$\frac{tn}{tn + fp}$	Calcula el proporción de la patrones negativos clasificado correctamente
Medida F (FM)	$\frac{2}{\frac{1}{tp} + \frac{1}{tn}}$	Describe el armonía entre recordar y valores de precisión

Donde tp representa el verdadero positivo, fp representa el falso positivo. mientras que tn representa el verdadero negativo y fn representa el falso negativo.

4.3 Aplicado Experimentos Y Evaluación

Resultados

Ejecutamos los algoritmos mencionados en el conjunto de datos. El conjunto de datos se dividió en 80% y 20% para entrenamiento y prueba, respectivamente. Comenzamos examinando los efectos del Doc2vec y luego del Word2vec en los diferentes algoritmos. Los algoritmos utilizados para crear modelos de aprendizaje y luego utilizarlos para predecir las etiquetas de los datos de prueba. Las tablas 2 y 3 muestran los resultados obtenidos de los diferentes clasificadores basados en Doc2vec y Word2vec.

Tabla 2. Resultados de Doc2vec

Evaluación	LR	RF	MLP	SVM	LSTM
Sensibilidad	0,946	0,876	0,941	0,948	0,624
Especificidad	0,948	0,942	0,951	0,961	0,565
exactitud	0,947	0,909	0,946	0,955	0,584
Puntuación F1	0,943	0,907	0,943	0,952	0,493

Tabla 3. Word2vec

Evaluación	LR	RF	MLP	SVM	LSTM
Sensibilidad	0,622	0,691	0,589	0,649	0,965
Especificidad	0,656	0,769	0,559	0,707	0,922
Precisión	0,639	0,724	0,570	0,674	0,943
Puntuación F1	0,634	0,743	0,511	0,694	0,941

Las cifras de 6, 7, 8, 9 y 10 representan los resultados de cada algoritmo individualmente basado en los dos modelos Doc2vec y Word2vec para mostrar los efectos de los dos métodos de extracción de características en los diferentes clasificadores.

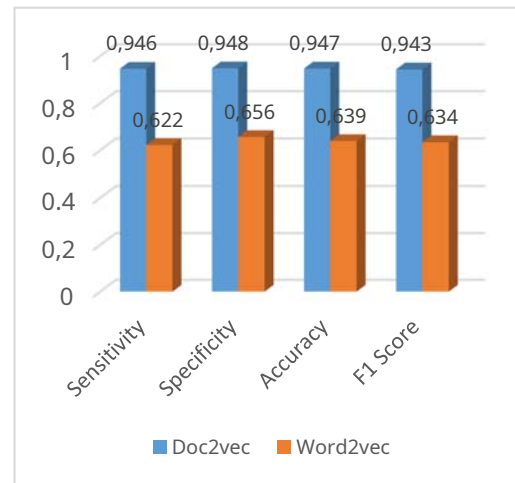


Figura 6. Resultados de LR

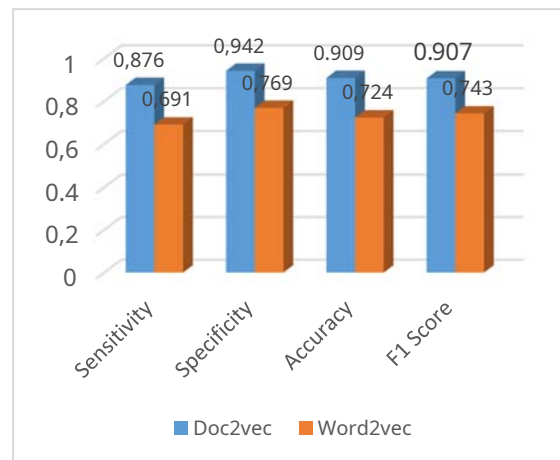


Figura 7. Resultados de RF

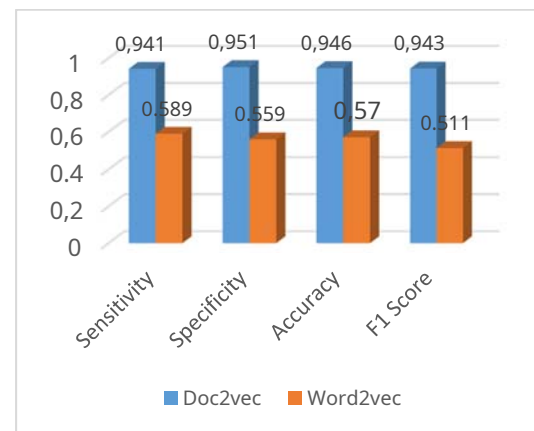


Figura 8. Resultados de MLP

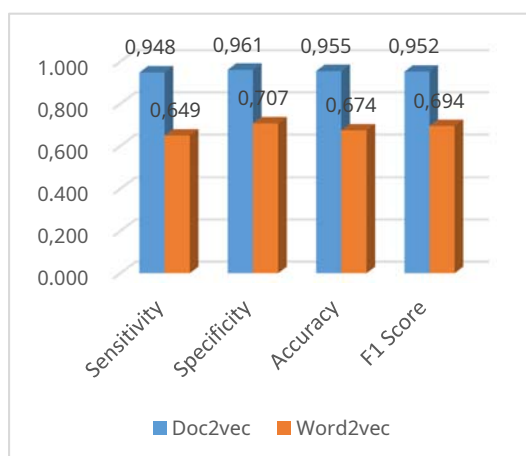


Figura 9. Resultados de SVM

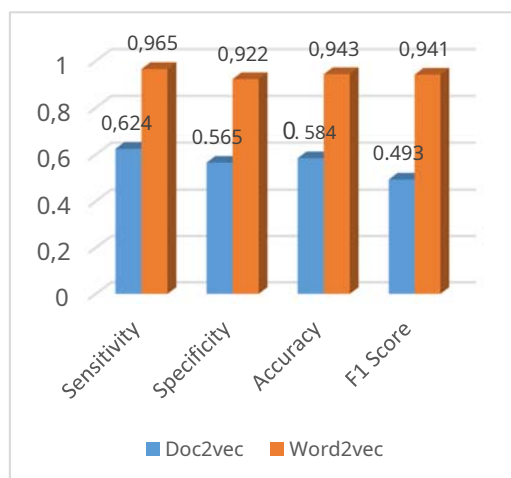


Figura 10. Resultados de LSTM

De los resultados anteriores, la incrustación de Doc2vec resulta una buena precisión en todos los algoritmos excepto LSTM, mientras que la incrustación de Word2vec produce una buena precisión solo con el clasificador LSTM.

Doc2Vec amplía Word2Vec proporcionando un vector de documento a la representación de la salida que incluye diversa información sobre el documento; también permite que el modelo aprenda algunos detalles sobre el orden de las palabras. La información guardada sobre el orden de las palabras en Doc2Vec es útil para SVM, LR, MLP y RF y los cuatro clasificadores se aprenden de ella y arrojan una precisión de clasificación aceptable y prometedora. La precisión resultante para SVM y LR es 59,5% y 94,7% respectivamente. Si bien la forma de mantener la información del pedido no da resultados aceptables con LSTM. Por lo tanto, también aplicamos Word2vec, y

ordenó la palabra más común en el conjunto de datos de entrenamiento (el número utilizado se basa en el total de palabras en el entrenamiento después de los pasos de preprocesamiento). Cada palabra se transfirió a un vector de 32 dimensiones, y luego cada vector de palabra se entrenó mediante la inserción de palabras en función de la similitud de palabras. Esta representación de Word2vec se alimentó a los diferentes clasificadores como una selección de características y la información de pedido se mantiene de esta manera, era más adecuada para LSTM y la precisión resultante fue del 94,3% y es una precisión prometedora.

Se aplicó otro experimento para comparar los clasificadores basados en el modelo Doc2vec con el trabajo en [27] en el mismo conjunto de datos, en su investigación aplicaron muchos clasificadores, y la Máquina de Vector de Soporte Lineal resultó en la mejor precisión del 92% mientras usaban TF-IDF como extracción de características. Aplicaron sus experimentos en un subconjunto de 2000 documentos del mismo conjunto de datos que se enfocan solo en la política. al comparar los resultados de los algoritmos SVM y LR basados en el modelo Doc2vec notamos que los resultados obtenidos superan a los resultados en

[27] como se muestra en la tabla 6.

Tabla 6 Comparación de Doc2vec con TF-IDF [27]

Evaluación	Doc2vec	TF-IDF
SVM	95,50%	92,00%
LR	94,70%	89%

En la tabla 6 mostramos los resultados obtenidos de Doc2vec con SVM y LR y los resultados obtenidos en [27] para los mismos clasificadores. Aunque concentraron su modelo en un tema específico. Los resultados de precisión obtenidos de los clasificadores que utilizan la extracción de características basada en Doc2vec supera los resultados de los mismos clasificadores basados en TF-IDF.

5. CONCLUSIÓN Y TRABAJOS FUTUROS

En esta investigación, describimos dos enfoques diferentes que se pueden utilizar como función métodos de extracción para el problema de detección de noticias falsas. El problema que se ha vuelto muy importante y sensible en nuestra vida social en los últimos años. Los dos enfoques propuestos son Word2vec y Doc2ve. El objetivo principal era aplicar Doc2vec y demostró con muchos clasificadores que se puede utilizar para entrenar a más de un clasificador y resultó una precisión muy prometedora con SVM y LR. Si bien los resultados de Doc2vec con LSTM no fueron aceptables. Entonces

introducimos Word2vec con una forma adecuada para que LSTM conservara la información de orden de las palabras y la precisión resultante fue del 94,3% que puede ser razonable y prometedor. Además, el modelo basado en Doc2vec comparado con otro modelo. El modelo basado en TF-IDF como extracción de características con N-gram. Los clasificadores SVM y LR que se basaron en el modelo Doc2vec resultaron en una mayor precisión que cuando se basaron en TF-IDF. En el futuro, los modelos se pueden aplicar en diferentes conjuntos de datos con más documentos, y la hibridación entre más de un clasificador puede mejorar la precisión de la clasificación.

REFERENCIAS:

- [1] Welbers, K. y M. Opgenhaffen, Redes sociales gatekeeping: un análisis de la influencia de gatekeeping de las páginas públicas de Facebook de los periódicos. *nuevos medios y sociedad*, 2018. 20 (12): p. 4728-4747.
- [2] Jankowski, NW, Investigación de noticias falsas: A examen selectivo de estudios empíricos. *Javnost-The Public*, 2018. 25 (1-2): pág. 248-255.
- [3] Carlson, M., Las noticias falsas como pánico moral informativo: la desviación simbólica de las redes sociales durante las elecciones presidenciales de Estados Unidos de 2016. *Información, comunicación y sociedad*, 2020. 23 (3): p. 374-388.
- [4] Sabeeh, VT, modelos de detección de noticias falsas asistidos por aprendizaje automático y conocimiento semántico. 2020, Universidad de Oakland.
- [5] Molina, MD, et al., "Fake news" no es simplemente información falsa: una explicación del concepto y taxonomía del contenido en línea. *Científico del comportamiento estadounidense*, 2019: p. 0002764219878224.
- [6] Dubey, AK, A. Singhal y S. Gupta, Sistema de detección de rumores mediante aprendizaje automático. *Revista Internacional de Investigación de Ingeniería y Tecnología (IRJET)* Vol. 7, N ° 5 (2020).
- [7] Lan, NPH, et al., Características fenotípicas y genotípicas de organismos productores de BLEE y AmpC asociados con bacteriemia en la ciudad de Ho Chi Minh, Vietnam. *Antimicrobiano Control de la resistencia y las infecciones*, 2017. 6 (1): p. 1-9.
- [8] Rubin, VL, et al. ¿Noticias falsas o verdad? utilizando señales satíricas para detectar noticias potencialmente engañosas. en *Actas del segundo taller sobre computacional enfoques para engaño detección*. 2016.
- [9] Tandoc Jr, EC, ZW Lim y R. Ling, definiendo "Noticias falsas" Una tipología de definiciones académicas. *Periodismo digital*, 2018. 6 (2): p. 137-153.
- [10] Fallis, D., ¿Qué es la desinformación? *Tendencias bibliotecarias*, 2015. 63 (3): p. 401-426.
- [11] Shu, K. y col. defender: Detección explicable de noticias falsas. en *las actas de la 25a Conferencia Internacional ACM SIGKDD sobre Descubrimiento de conocimiento y minería de datos*. 2019.
- [12] Meel, P. y DK Vishwakarma, noticias falsas, rumores, contaminación de la información en las redes sociales y la web: una encuesta contemporánea sobre el estado del arte, los desafíos y las oportunidades. *Sistemas expertos con aplicaciones*, 2019: p. 112986.
- [13] Ozbay, FA y B. Alatas, detección de noticias falsas en las redes sociales en línea mediante algoritmos de inteligencia artificial supervisados. *Physica A: Mecánica estadística y sus aplicaciones*, 2020. 540: pág. 123174.
- [14] Collins, B., et al., Tendencias en la lucha contra la falsificación noticias en las redes sociales: una encuesta. *Revista de Información y Telecomunicaciones*, 2020: p. 1-20.
- [15] Mackey, TK y G. Nayyar, Una revisión de las tecnologías digitales existentes y emergentes para combatir el comercio mundial de medicamentos falsos. *Opinión de experto sobre seguridad de los medicamentos*, 2017. 16 (5): p. 587-602.
- [16] Halder, M., M. Suresh y S. Rao, Impact of Discurso en aplicaciones modernas de PNL basadas en IA. *Smitha Rao, Impacto del discurso en aplicaciones modernas de PNL basadas en IA (31 de julio de 2020)*. Instituto de Académicos (InSc), 2020.
- [17] Wang, WY, "mentiroso, mentiroso en llamas": un nuevo conjunto de datos de referencia para la detección de noticias falsas. preimpresión de arXiv arXiv: 1705.00648, 2017.
- [18] Gupta, H. y col. Un marco para tiempo real detección de spam en Twitter. en *2018 10th International Conference on Communication Systems & Networks (COMSNETS)*. 2018. IEEE.
- [19] Al-Ansari, K., Encuesta sobre técnicas de incrustación de palabras en el procesamiento del lenguaje natural.
- [20] Le, Q. y T. Mikolov. Repartido representaciones de oraciones y documentos. en *conferencia internacional sobre aprendizaje automático*. 2014. PMLR.
- [21] Pranckevičius, T. y V. Marcinkevičius, Comparación de bayes ingenuos, bosque aleatorio, árbol de decisión, máquinas de vectores de apoyo y clasificadores de regresión logística para la clasificación de revisiones de texto. *Revista Báltica de Computación Moderna*, 2017. 5 (2): p. 221.

- [22] Crabbé, A., et al., Clasificador MLP de red neuronal. 2020.
- [23] Demotte, P. y col. Análisis de sentimiento de los comentarios de noticias cingaleses utilizando redes LSTM de estado de oración. en 2020 Moratuwa Jornada de Investigación en Ingeniería (MERCon). 2020. IEEE.
- [24] LIN, H., L. LI y H. WANG, Encuesta sobre investigación y aplicación de máquinas de vectores de apoyo en sistemas de transporte inteligentes. Revista de fronteras de la ciencia y la tecnología de la computación, 2020. 14 (6): p. 901-917.
- [25] Alam, S., et al., Clasificadores de vectores de soporte de una clase: una encuesta. Sistemas basados en el conocimiento, 2020. 196: p. 105754.
- [26] Ahmed, H., I. Traore y S. Saad, Detección de mensajes no deseados de opinión y noticias falsas mediante la clasificación de texto. Seguridad y privacidad, 2018. 1 (1): p. e9.
- [27] Ahmed, H., I. Traore y S. Saad. Detección de noticias falsas en línea mediante análisis de n-gramas y técnicas de aprendizaje automático. en conferencia internacional sobre sistemas inteligentes, seguros y confiables en entornos distribuidos y en la nube. 2017. Springer.