

convirtiéndose en un tema típico entre los adolescentes en estos días [3]. Estos incorporan la difusión de chismes sobre un individuo, peligros y comportamiento inapropiado. El acoso cibernético influye desfavorablemente en la persona en cuestión y provoca una variedad de reacciones apasionadas, por ejemplo, redujo la confianza, expandió las consideraciones autodestructivas, indignación y desdicha [4]. Los jóvenes son víctimas de estos ataques porque no logran captar la sofistería y la conducta ensimismada del agresor. Otro territorio donde una ubicación engañosa es de importancia central es con el mayor número de historias falsas, también conocido como noticias falsas, en Internet. Informes tardíos recomiendan que el resultado de las elecciones presidenciales de EE. UU. Se debe al aumento de noticias falsas en línea [5]. Los defensores utilizan argumentos que, aunque de vez en cuando persuaden, no son realmente legítimos. La vida basada en la web, por ejemplo, Facebook y Twitter, se han convertido en los propulsores de esta publicidad política con propósito. Las naciones de todo el mundo, por ejemplo, Francia, están utilizando estrategias que mantendrían la difusión de noticias falsas en medio de sus decisiones [5]. A pesar de que estas medidas pueden ayudar, existe un requisito apremiante para que la red fonética computacional diseñe estrategias productivas para luchar contra FakeNews dado que las personas no distinguen el doble trato. Por lo tanto, se necesita extraordinariamente un indicador programado para aliviar los efectos negativos genuinos provocados por las noticias falsas [6]. Hay muchas metodologías, como algoritmos de seguimiento basados en filtros de correlación [7], algoritmo de mínimos cuadrados no negativos [8], Método de selección de muestras representativas en línea [9], marco de regularización [10], modelo fusionado de múltiples funciones [11] ha sido introducido.

Todo el trabajo se presenta en cuatro secciones de la siguiente manera. El segmento 2 describe los trabajos relacionados en el campo de la detección de noticias falsas. En el segmento 3 se habla de una revisión sobre las clasificaciones de verosimilitud y el algoritmo. El segmento 4 presenta los resultados experimentales. En esta parte, también se presentan más discusiones y análisis. El segmento 5 es el breve resumen de este trabajo y el plano de los trabajos futuros.

2. Revisión de la literatura

Hay numerosos recados identificados con el reconocimiento de noticias falsas, por ejemplo, descubrimiento de rumores [12] y detección de spam [13]. Siguiendo el trabajo anterior [14, 15], indicamos el significado de noticias falsas como noticias creadas deliberadamente y que pueden confirmarse como falsas. En la tarea de identificación de noticias falsas, el principal desafío son los medios para reconocer las noticias según lo indicado por los aspectos más destacados. Los aspectos más destacados se pueden extraer de publicaciones, entornos sociales e incluso imágenes adjuntas.

El reconocimiento de noticias falsas ha surgido a la luz de una preocupación legítima por los especialistas en los últimos tiempos y se han propuesto algunas metodologías. En los últimos tiempos, se ha intentado utilizar el contenido de las noticias para la empresa de identificación [dieciséis, 17]. Wang [18] utiliza CNN para ordenar el contenido de noticias falsas. Shu y col. [dieciséis, 17] utilizar la instalación de sustancia inactiva del registro como uno de los aspectos más destacados para la empresa de reconocimiento. Hay algunos trabajos diferentes que hacen uso del contenido sustancial.

Ruchansky y col. [19] utilizar el compromiso social a nivel de puesto para captar las distinciones en los diseños de compromiso transitorio entre noticias falsas y genuinas. Dado que las personas expresan sus sentimientos hacia las noticias a través de publicaciones de redes sociales basadas en la web, es sensato utilizar publicaciones de vida basadas en la web como un componente potencial para la ubicación destacada. Shu y col. [20, 21] utilizan diferentes aspectos destacados del cliente que se relacionan con los artículos de noticias para reconocer las noticias falsas.

En este artículo, trabajaremos en los algoritmos ML más utilizados para elegir los mejores algoritmos de clasificación. Entre ellos, los tres con mejor desempeño se utilizarán en el Clasificador de votación. A partir de esto, podemos obtener el mejor resultado exacto de estos clasificadores. Nuestro trabajo proporciona resultados más precisos en comparación con los demás trabajos.

3. Metodología

La arquitectura propuesta se puede dividir en muchas subsecciones. El diagrama de flujo del trabajo se puede ver en la Fig. 1.

3.1 Recolección de datos

En primer lugar, se necesita un conjunto de datos con noticias reales y falsas. El sistema propuesto se prueba en el conjunto de datos de 6500 datos de los cuales aproximadamente 3252 datos son falsos y 3259 datos son reales que ha sido utilizado por Wang [dieciséis]. Es un conjunto de datos de combinación de noticias reales y falsas.

3.2 Preprocesamiento

En todos los índices de datos de actualidad, que contienen varios errores, se actualizan y eliminan para obtener resultados definitivos en el índice de datos. En esta recopilación de datos de movimiento, se cambia y se compone en un plan legítimo antes de que los clasificadores se asocien en el índice de datos. El registro se ha cuidado adecuadamente antes de que se le asocien clasificadores. El conjunto de datos se basa principalmente en el idioma inglés. Para el preprocesamiento se aplica la técnica de procesamiento de lenguaje natural de datos (NLP) donde solo tomamos palabras en inglés. Ayuda a mejorar la precisión. Después de eso, se realizan la transformación de texto y los archivos binarios en el conjunto de datos para facilitar el preprocesamiento de datos.

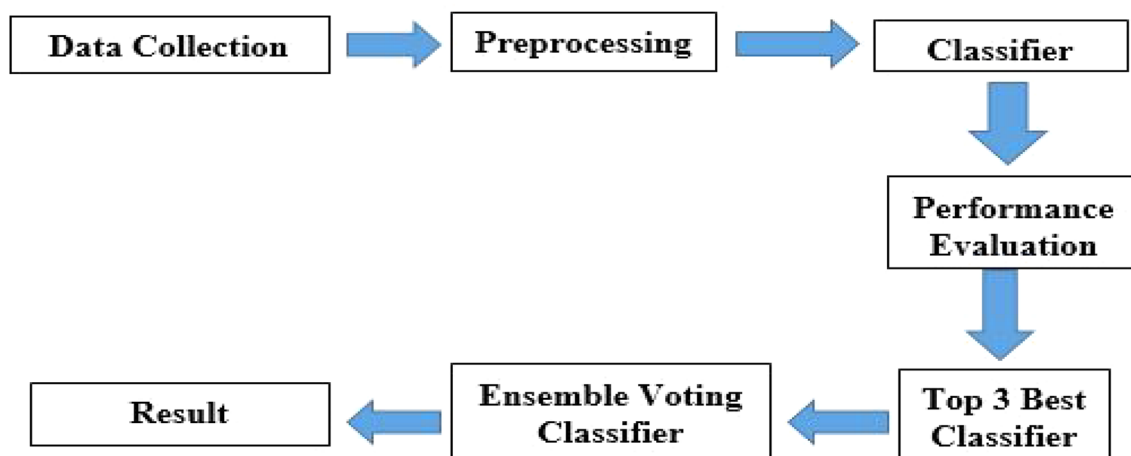


Figura 1 Diagrama de flujo de la arquitectura de conjunto

3.3 Clasificador

Después de tener el documento preprocesado, todos los clasificadores conocidos, en particular, Vector de soporte, Vecinos más cercanos K, Ada Boost, Naïve Bayesian, Neural Network, DecisionTree, Gradient Boosting, Extreme Gradient Boosting, Random Forest, Logistic Regression, etc. se ha aplicado para descubrir incluye en función de qué falsificación se detecta.

3.4 Evaluación del desempeño

Para la evaluación del desempeño, se ha utilizado la técnica de validación cruzada. Aquí se ha realizado la validación cruzada de K-fold. En consecuencia, el conjunto de datos se divide en 10 K veces [22].

En la implementación de este paso, se ha utilizado la función `model_selection` de `scikit-learn`. La subfunción `estratificada K-Fold` se ha utilizado para dividir el conjunto de datos de entrenamiento en K-fold para la validación cruzada, la subfunción `cross_val_score` se ha utilizado para observar las puntuaciones de validación cruzada de los clasificadores ML y la subfunción `GridSearchCV` se ha utilizado para hiper- sintonizar los clasificadores ML.

Como resultado de la aplicación de todos los clasificadores, todos se encuestaron en función de las estimaciones de ejecución, como la puntuación de la prueba, la puntuación de ROC, la puntuación de precisión, el valor de recuperación, etc., para comprender el mejor clasificador.

3.5 Elección de los 3 mejores clasificadores

Después de la evaluación del desempeño de los diversos clasificadores ML tradicionales usados, se han reconocido los tres mejores clasificadores. En ese momento, estos tres principales

Los clasificadores se utilizarán en el siguiente paso para ajustar para obtener el mejor resultado del conjunto de datos. Entonces se utilizará el clasificador de votación.

3.6 Utilización del clasificador de votaciones por conjuntos

Para el clasificador de conjunto, aquí este artículo hablará sobre el clasificador de votación. Los tres clasificadores principales se utilizarán para esta clasificación de votación para obtener la mejor ejecución y resultado.

3.7 Resultados

En el último avance, la presentación del Clasificador de votaciones será encuestada en función de las estimaciones de ejecución, como la puntuación de la prueba, la puntuación de ROC, la puntuación de precisión, el valor de recuperación, etc. Los resultados en ese punto serán contrastados y otros trabajos importantes para evaluar los resultados.

3.8 Clasificador de votación por conjunto

El clasificador de votación por conjunto [23, 24] es un metaclasificador para unir clasificadores de aprendizaje automático similares o hábilmente sin precedentes para la clasificación y detección. El clasificador de votaciones por conjuntos ejecuta votaciones "duras" y "blandas".

3.9 Voto en duro

En el conjunto difícil, la votación es la instancia más sencilla de la mayor parte de la votación. Aquí, determinaríamos la marca de clase Y mediante la votación de la parte del león de cada clasificador C:

$$= \text{modo } \{C1(x), C2(x), \dots, C_m(x)\}$$

3.10 Votación blanda

En la votación de conjunto suave, visualizamos los nombres de las clases sujetos a las probabilidades previstas p para el clasificador, este procedimiento quizás se recomienda si los clasificadores son particularmente ajustado.

$$= \underset{(j=1)}{\operatorname{argmax}_i} \sum_{j=1}^{\text{metro}} W_j \text{PAG}_{ij} \in \{.,\}, \quad [j = 1, 2, \dots, \text{metro}]$$

donde W_j es el montón que se puede distribuir al j -ésimo clasificador.

Las breves discusiones de algunos clasificadores que se han utilizado para seleccionar los tres clasificadores principales para el conjunto se encuentran a continuación:

3.11 Clasificador bayesiano ingenuo

Fue uno de los principales procedimientos de caracterización utilizados para la detección de noticias falsas. Se necesita una oportunidad en el Teorema de Bayes de la probabilidad para comprobar si las noticias que se acercan son falsas o no. Inicialmente, el filtro de este clasificador debe estar preparado para buscar noticias falsas. Preparar una recopilación de datos implica que el canal recibe muchas palabras que el cliente da al reconocer físicamente la noticia como falsa o no. A través de la contribución de un cliente, el clasificador está actualmente preparado y puede navegar por las noticias que se acercan. Este clasificador comprueba la verosimilitud de las palabras del set de preparación en las noticias que se aproximan y con el resultado adquirido puede canalizar noticias falsas. Hace un organizador alternativo para la falsificación y mueve esas noticias directamente a la carpeta. A pesar de que es completamente antiguo, [25].

3.12 K-Vecino más cercano

En este procedimiento de clasificación, un conjunto de formación contiene un ejemplo de noticias que puede ayudar a distinguir si las noticias que se acercan son falsas o no. La noticia que se acerca se contrasta con el conjunto de entrenamiento con el descubrimiento de su parte k más cercana al contrastarlo y el conjunto de entrenamiento y su archivo más comparable se encuentra y luego se reconoce como falso, a la luz de qué grupo se encontró su informe más k comparativo [26].

3.13 Clasificador de máquina de vectores de soporte

En esta técnica, se forma un plano de elección para aislar noticias falsas y genuinas. Están aislados por un límite de elección que tiene ciertas condiciones para aislar esta noticia. A

El conjunto de formación se enmarca para la organización y la próxima noticia se contrasta y el conjunto de preparación. Como k -más cercano, vecino, las noticias que se acercan se contrastan con el conjunto de preparación con una similitud descubierta entre las noticias que se acercan y el conjunto de entrenamiento. Un trabajo parcial, K se utiliza para decidir la cercanía y dependiendo de esto las noticias se ordenan en el plano de elección [27].

3.14 Red neuronal artificial

La red neuronal artificial actúa como una mente humana falsa. Las neuronas falsas están interconectadas para enmarcar este sistema y la información se transmite a través de él para su aprendizaje. Al igual que la mente humana, aprende por los precedentes y, en medio de la preparación, la información pasa por el sistema para que pueda aprender y ajustarse según los modelos [28]. Cambian su estructura dependiendo de los datos de los modelos con el objetivo de que se pueda dar forma a una disposición superior de agrupación [29].

3.15 Clasificador de árbol de decisión

Decision Tree es un enfoque de visualización profético que se utiliza en aprendizaje automático, minería de datos y conocimientos. Realiza un modelo basado en unos pocos factores de información que predice la estimación de una variable objetiva. Es un cálculo ampliamente utilizado que sigue la metodología voraz en cada división y fabrica dinámicamente un árbol. Cada eje de un árbol de decisión habla del análisis de la calidad, las ramas hablan de las secuelas de los exámenes y los ejes de las hojas contienen las marcas de clase. Las partes del árbol de elección se seleccionan con el objetivo final de que limiten la contaminación y amplíen la inmaculación del subconjunto que se está desarrollando [30].

3.16 clasificador RandomForest

También es una clasificación basada en árboles. Un arreglo se realiza mediante la creación de varios árboles de decisión distintivos, cada uno de los cuales tiene una estructura de componentes alternativa. A partir de ese momento, se designa una clase en función de la mayor parte de los votos de los árboles distintivos [17].

3.17 clasificador Ada Boosting

Un clasificador Ada Boost es un meta predictor que comienza ajustando un clasificador en el primer conjunto de datos y luego ajusta duplicados adicionales del clasificador en el conjunto de datos equivalente, pero donde una gran cantidad de ejemplos agrupados de manera inexacta se equilibran con el objetivo final de que los clasificadores consecuentes se centren más en casos problemáticos. [31].

3.18 Clasificador de aumento de gradiente

Gradient Boosting es un sistema de aprendizaje de máquina para problemas de recaída y clasificación, que crea un programa de pronóstico como una recopilación de modelos de expectativas débiles, normalmente DecisionTrees. Fabrica el modelo en una fase de diseño astuto como lo hacen otras estrategias de impulso, y las resume al permitir la mejora de un trabajo de desgracia diferenciable discrecional [32].

3.19 Clasificador de regresión logística

La regresión logística consiste en evaluar los parámetros de un modelo logístico; es un tipo de recaída binomial. Científicamente, un modelo calculado doble tiene una variable dependiente con dos

cualidades concebibles, por ejemplo, pasa / no pasa, gana / pierde, vivo / muerto; a estos se les habla mediante una variable de puntero, donde las dos cualidades están marcadas como "0" y "1". En el modelo logístico, las posibilidades logarítmicas de la estimación denominada "1" es una combinación directa de al menos un factor autónomo; cada uno de los factores libres puede ser un factor doble o una variable constante. La probabilidad relativa del "1" marcado estimado puede diferir en algún lugar del rango de 0 y 1, de ahora en adelante la denominación; la capacidad que cambia a lo largo de las posibilidades logarítmicas a la probabilidad es la capacidad estratégica, por lo que el nombre [33].

4 Análisis de resultados

A partir del conjunto de datos recopilados, se incluyeron alrededor de 6500 noticias, de las cuales alrededor de 3252 noticias son falsas y 3259 noticias son reales, que se han referido en la recopilación de datos de la metodología. La representación del gráfico circular de este conjunto de datos se ha proporcionado en la Fig.2. El conjunto de datos que se utiliza en este trabajo se verifica y analiza con once técnicas diferentes de algoritmos de clasificación de Machine Learning que se utilizan en la validación cruzada, que son las siguientes: (a) Random Forest, (b) Ada Boosting, (c) Gradient Boosting, (d) Árboles adicionales, (e) Regresión logística, (f) K-Vecinos, (g) Árbol de decisión, (h) Bayes ingenuo multinomial, (i) Percepción multicapa, (MLP), (j) Máquina de vectores de soporte y (k) Incremento de gradiente extremo. Después de preprocesar el conjunto de datos, como limpiar los valores faltantes, transformar los datos en vectores, etc., los datos de entrenamiento se dividen en diez. Después de eso, medimos la puntuación de validación cruzada de estos once clasificadores. Figura3 y mesa 1 representa la puntuación de validación cruzada de estos once clasificadores.

Después del resultado de las puntuaciones de validación cruzada, elegimos los tres mejores algoritmos de clasificación de ML (1) MLP, (2)

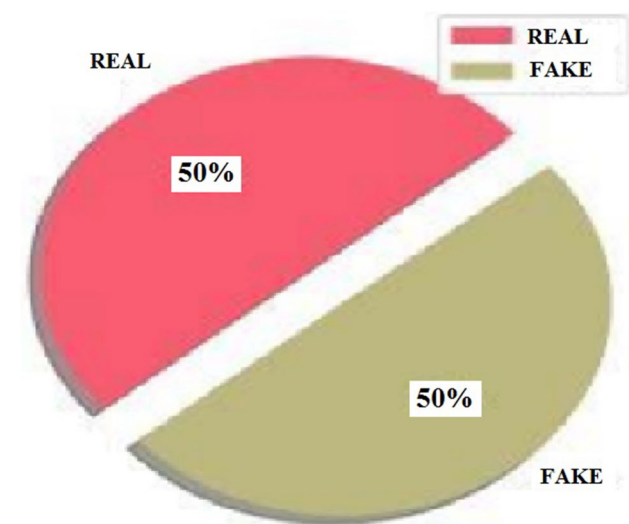


Figura 2 La representación de gráfico circular del conjunto de datos

Fig. 3 Representación de la puntuación de validación cruzada de varios clasificadores de AA mediante el uso de un gráfico de barras

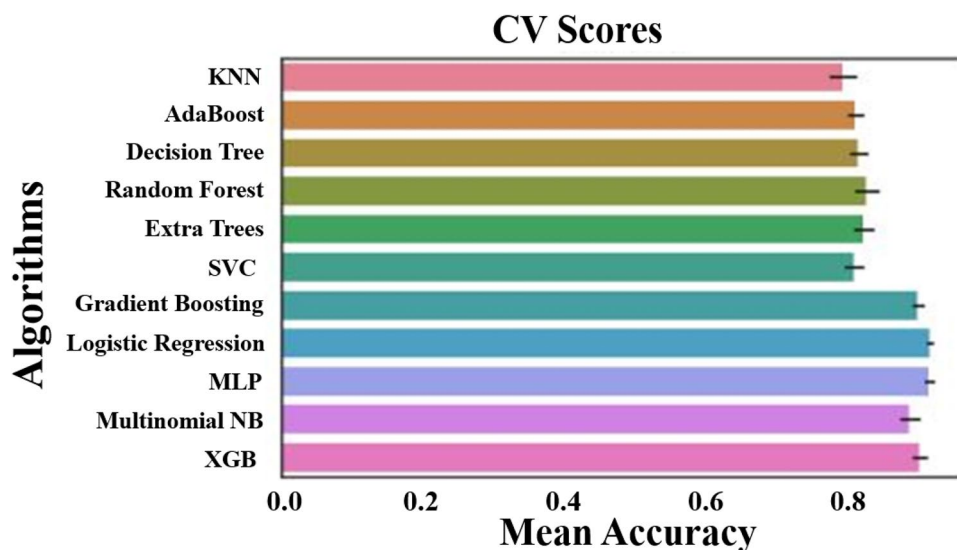


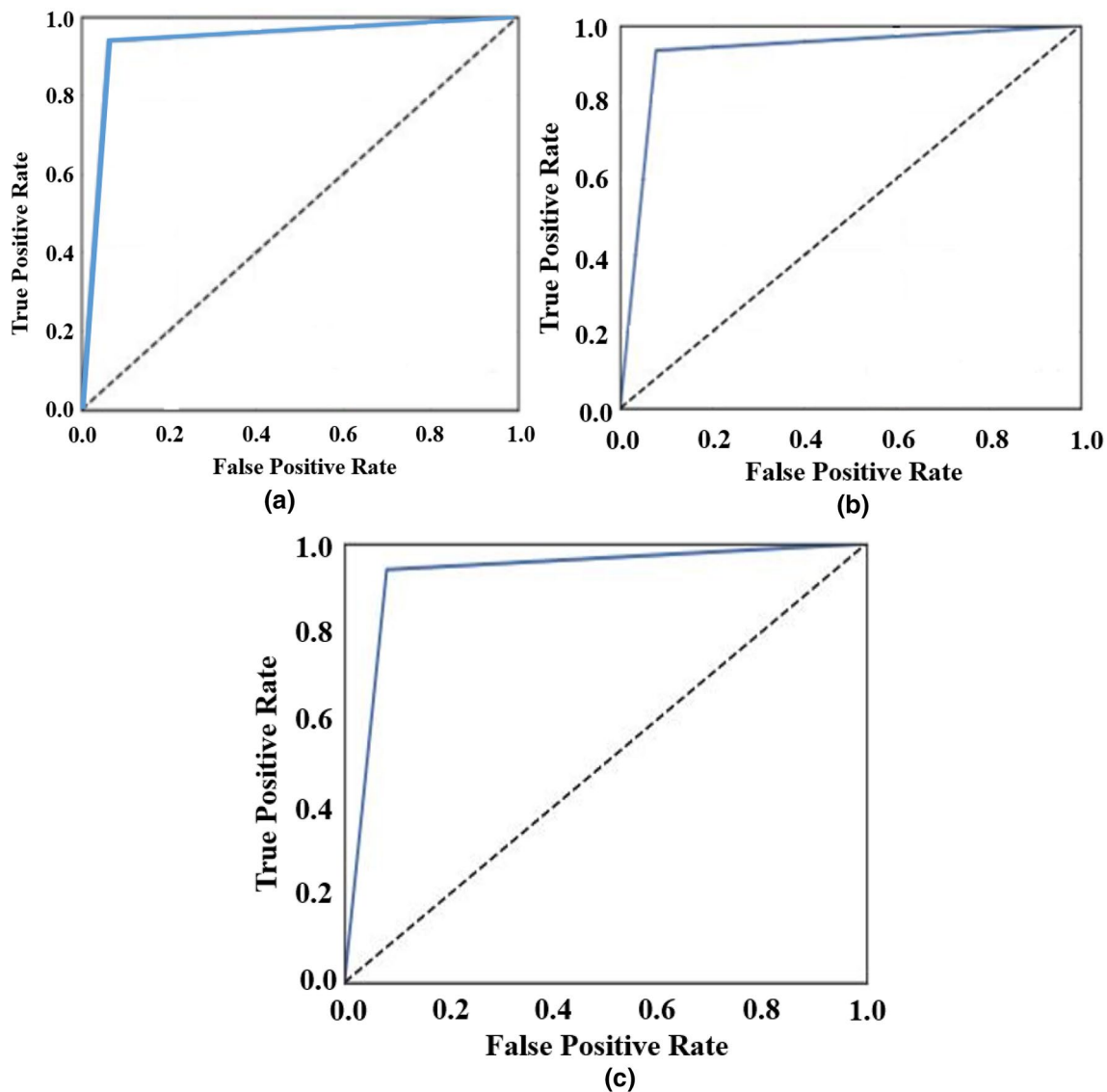
tabla 1 Puntuación de validación cruzada de varios clasificadores de AA

Tipo de clasificación	Validación cruzada puntuaje de ción (%)
Vecinos K	79,63
Ada Boost	82,62
Árbol de decisión	82,69
Bosque aleatorio	83,59
Árbol extra	82,48
SVC	84,75
Aumento de gradiente	90,30
Regresión logística	91,97
Percepción multicapa (MLP)	91,9
Multinomial Naïve Bayes X-	89,17
Gradient Boosting	90,53

Regresión logística y (3) aumento de gradiente X para el siguiente paso. Aquí, hiper-sintonizaremos estos tres clasificadores para obtener los mejores resultados de ellos (Fig.4).

Para la clasificación MLP, ajustamos los parámetros en función de alfa, tamaño de capa oculta, iteraciones máximas, solucionador y estado aleatorio. Después de ajustar el parámetro, obtenemos mejores resultados para estos parámetros: 'alfa': 0.01, 'tamaños de capa oculta': 14, 'iteración máxima': 100, 'estado aleatorio': 0, 'solucionador': 'lbfgs'. Para MLP, la mejor puntuación 92,59, precisión 93,83, precisión 85,94, recuperación 85,49 y puntuación ROC 93,83 (Fig. 5).

Para la clasificación de X-Gradient Boosting, ajustamos los parámetros en función de gamma, tasa de aprendizaje, pérdida, profundidad máxima, hoja de muestra mínima y n_estimators. Después de ajustar el parámetro, obtenemos mejores resultados para estos

**Figura 4** Curva ROC de **a** MLP, **B** Aumento de gradiente X y **C** Regresión logística

parámetros: 'gamma': 1, 'tasa de aprendizaje': 0.1, 'pérdida': 'desviación', 'profundidad máxima': 15, 'hojas mínimas de muestras': 5, 'n_estimators': 100. Para X-Gradient Boosting, la mejor puntuación 92,92, precisión 92,87, precisión 86,67, recuperación 87,67 y puntuación ROC 92,87.

Para la clasificación de regresión logística, ajustamos los parámetros en función de la tolerancia, la iteración máxima, C, la escala de intercepción, la penalización y el solucionador. Después de ajustar el parámetro, obtenemos mejores resultados para estos parámetros: 'C': 0.1, 'escala de intercepción': 1, 'iteración máxima': 100, 'penalización': 'l2', 'solucionador': 'liblinear', 'tolerancia': 0.0001. Para Regresión logística, la mejor puntuación 92,57, precisión 93,03, precisión 85,44, recuerdo 90,01 y puntuación ROC 93,03 (tabla 2).

Después de recibir los mejores resultados de estos tres algoritmos de clasificación ML, los usaremos en el clasificador de votación para obtener una puntuación máxima en la prueba. El puntaje final de la prueba del clasificador de votación blanda Ensemble es 94,47. El otro

Los parámetros de este clasificador son Precision 95, Recall 95, F1 95 y ROC score 94.49. Por otro lado, el puntaje final de la prueba del clasificador de votación dura por conjuntos es 93,99. Los otros parámetros de este clasificador son Precision 94, Recall 94, F1 94 y puntuación ROC 93,98 (tabla 3).

A partir del análisis, queda claro que el clasificador de votación por conjuntos proporciona la mejor puntuación de prueba de todos los algoritmos de clasificación de ML individuales.

5. Conclusiones

En este trabajo de investigación, se propone una técnica única de clasificador de votación por conjuntos basada en múltiples clasificadores para decidir noticias tanto reales como falsas. Varios algoritmos de clasificación de aprendizaje automático tradicionales y más utilizados se han utilizado para el conjunto de datos dado de noticias para

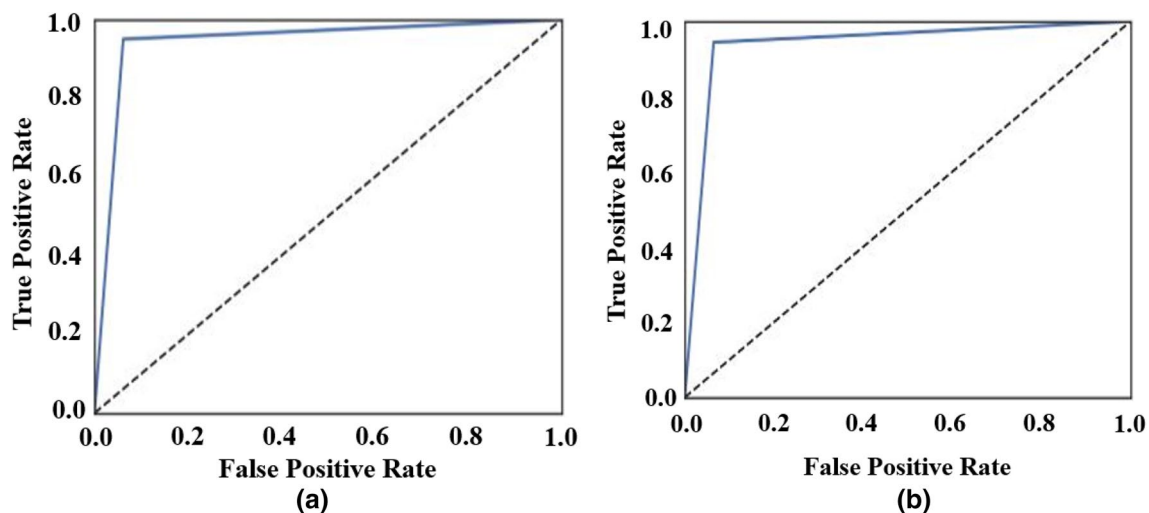


Figura 5 Curva ROC del clasificador de votación **a** voto suave y **B** voto duro

Tabla 2 Representación de la mejor puntuación, exactitud, precisión, recuperación y puntuación ROC después de ajustar MLP, X-Gradient Impulso y Logística Regresión

Tipo de clasificación	Mejor puntuación	Precisión	Precisión	Recordar	Puntuación ROC
MLP	92,59	93,83	85,94	85,49	93,83
Aumento de gradiente X	92,92	92,87	86,67	87,67	92,87
Regresión logística	93.03	98,21	85,44	90.01	93.03

Tabla 3 Informe para clasificador de votación por conjunto

Tipo de votación	Escribe	Resultado de la prueba	Precisión	Recordar	F1	Puntuación ROC
Suave	Verdadero	94,47	95	94	94	94,49
	Falso		94	95	95	
	Promedio		95	95	95	
Duro	Verdadero	93,99	94	94	94	93,98
	Falso		94	94	94	
	Promedio		94	94	94	

planearlos para convertirlos en reales y falsos. Los resultados mostraron que este elemento sería astuto de usar con respecto a la exactitud, precisión, recuerdo, puntuación ROC, F1 para controlar utilizando agregaciones de noticias y técnicas de clasificación similares. Además, los resultados mostraron que el clasificador de votación por conjunto demostró mejores puntajes de suficiencia cuando se destacó de los resultados obtenidos por los clasificadores individuales.

Hay varias opciones interesantes para trabajos futuros. Una es hacer uso de otras funciones disponibles en el conjunto de datos, como retweets, redes sociales, comunidad de imágenes de Instagram y conocer los aspectos más destacados del descubrimiento de noticias falsas. Del mismo modo, nuestra estructura propuesta podría extenderse para identificar noticias falsas continuamente a medida que se actualizan en forma de transmisión. También se pueden implementar técnicas de aprendizaje profundo para mejorar la precisión y los puntajes de las pruebas.

6 Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Fondos Este estudio fue autofinanciado.

Referencias

- Bond CF Jr, DePaulo BM (2006) Precisión de los juicios de engaño. *Personal Soc Psychol Rev* 10 (3): 214–234
- DePaulo BM, Lindsay JJ, Malone BE, Muhlenbruck L, Charlton K, Cooper H (2003) Señales para el engaño. *Psychol Bull* 129: 74
- Smith PK, Mahdavi J, Carvalho M, Fisher S, Russell S, Tippett N (2008) Cyberbullying: su naturaleza e impacto en alumnos de secundaria. *J Child Psychol Psychiatry* 49: 376–385
- Kowalski RM, Giumetti GW (2017) Bullying en la era digital. En: Martellozzo E, Jane EA (eds) *El ciberdelito y sus víctimas*. Routledge, Abingdon, págs. 167–186
- Krishnamurthy G, Majumder N, Poria S, Cambria E (2018) Un enfoque de aprendizaje profundo para la detección de engaños multimodal. Preimpresión de ArXiv arXiv: 1803.00344
- Wang Y, Ma F, Jin Z, Yuan Y, Xun G, Jha K, Su L, Gao J (2018) EANN: redes neuronales adversas de eventos para la detección multimodal de noticias falsas. En: *Actas de la 24ª conferencia internacional ACM SIGKDD sobre descubrimiento de conocimientos y minería de datos*, págs. 849–857. ACM
- Yuan D, Lu X, Li D, Liang Y, Zhang X (2019) Redetección del filtro de partículas para seguimiento visual a través de filtros de correlación. *Multimed Tools Appl* 78 (11): 14277–14301
- Ou W, Yuan D, Liu Q, Cao Y (2018) Seguimiento de objetos basado en la selección de muestras representativas en línea a través de mínimos cuadrados no negativos. *Multimed Tools Appl* 77 (9): 10569–10587
- Ou W, Yuan D, Li D, Liu B, Xia D, Zeng W (2017) Seguimiento visual basado en parches con selección de muestras representativas en línea. *Imagen de electrones J* 26 (3): 033006
- Yuan D, Lu S, Li D, Zhang X (2019) Refinación de gráficos mediante un marco de regularización iterativo. *SN Appl Sci* 1 (5): 387
- Yuan D, Zhang X, Liu J, Li D (2019) Un modelo fusionado de múltiples funciones para el seguimiento de objetos visuales mediante filtros de correlación. *Aplicación de herramientas mul-timed* 78 (19): 27271–27290
- Jin Z, Cao J, Jiang YG, Zhang Y (2014) Evaluación de la credibilidad de las noticias en microblogs con un modelo de propagación jerárquico. En: *2014 IEEE international conference on data mining (ICDM)*. IEEE, págs. 230–239
- Shen H, Ma F, Zhang X, Zong L, Liu X, Liang W (2017) Descubriendo spammers sociales desde múltiples puntos de vista. *Neurocomputing* 225 (2017): 49–57
- Ahmed H, Traore I, Saad S (2017) Detección de noticias falsas en línea mediante el análisis de N-gram y técnicas de aprendizaje automático. En: *Conferencia internacional sobre sistemas inteligentes, seguros y confiables en entornos distribuidos y en la nube*. Springer, Cham
- Shu K, Sliva A, Wang S, Tang J, Liu H (2017) Detección de noticias falsas en las redes sociales: una perspectiva de la minería de datos. *ACM SIGKDD Explor Newsl* 19 (1): 22–36
- Shu K, Wang S, Liu H (2017) Explotación de la triple relación para la detección de noticias falsas. *ArXiv*: 1712.07709
- Gilda S (2017) Evaluación de algoritmos de aprendizaje automático para la detección de noticias falsas. En: *2017 IEEE 15th student conference on research and development (SCORED)*. IEEE
- Wang WY (2017) "Mentiroso, mentiroso en llamas": un nuevo conjunto de datos de referencia para la detección de noticias falsas. *ArXiv*: 1705.00648
- Ruchansky N, Seo S, Liu Y (2017) CSI: un modelo profundo híbrido para la detección de noticias falsas. En: *Actas de la ACM de 2017 sobre la conferencia sobre gestión de la información y el conocimiento*. ACM, págs. 797–806
- Shu K, Bernard HR, Liu H (2018) Estudiar noticias falsas a través del análisis de redes: detección y mitigación. *ArXiv*: 1804.10233
- Shu K, Wang S, Liu H (2018) Comprensión de los perfiles de usuario en las redes sociales para la detección de noticias falsas. En: *conferencia IEEE sobre procesamiento y recuperación de información multimedia (MIPR)*, págs. 430–435
- Mahabub A (2019) Un enfoque de votación sólido para la predicción de la diabetes utilizando técnicas tradicionales de aprendizaje automático. *SN Appl Sci* 1 (12): 1667
- Mahabub A, Mahmud MI, Hossain MF (2019) Un sistema robusto para el filtrado de mensajes que utiliza un enfoque supervisado de aprendizaje automático conjunto. *ICIC Express Lett Part B Appl* 10: 805–811. <https://doi.org/10.24507/iceib.10.09.805>
- Mahabub A, Habib AZSB (2019) Un enfoque de votación de clasificación de modulación para redes inalámbricas. En: *Actas de la 6ª conferencia internacional sobre redes, sistemas y seguridad*, págs. 133–138. ACM
- Granik M, Mesyura V (2017) Detección de noticias falsas utilizando un clasificador de Bayes ingenuo. En: *2017 IEEE primera conferencia de Ucrania sobre ingeniería eléctrica e informática (UKRCON)*. IEEE
- Tretyakov K (2004) Técnicas de aprendizaje automático en el filtrado de spam. *Instituto de Ciencias de la Computación, Universidad de Tartu, Tartu*
- Iyer A, Pandey A, Pamnani D, Pathak K, Hajgude J (2014) "Filtrado y análisis de correo electrónico mediante algoritmos de clasificación", Departamento de TI, VESIT. *Int J Comput Sci Issues* 11 (4): 1
- Patidar V, Singh D, Singh A (2013) Una técnica novedosa de clasificación de correo electrónico para la detección de spam. *Int J Appl Inf Syst* 5 (10): 15–19
- Annasaheb AB, Verma VK (2016) Técnicas de clasificación de minería de datos: una encuesta reciente. *Int J Emerg Technol Eng Res* 4 (8): 51–54
- Shrivastava S, Anju R (2017) Detección de correo no deseado mediante técnicas de minería de datos. En: *2017 Conferencia internacional sobre comunicación inteligente y técnicas computacionales (ICCT)*, págs. 61–64. IEEE
- Hastie T, Rosset S, Zhu J, Zou H (2009) Multi-class AdaBoost. *Interfaz estadística* 2 (3): 349–360

32. Parviz M, Shahram Moin M (2008) Estimación de polinomios multivariante basada en el aumento de gradiente en biometría multimodal. En: Conferencia internacional sobre computación inteligente, págs. 471–477. Springer, Berlín
33. Harrell FE (2015) Estrategias de modelado de regresión. Serie Springer en estadística, 2ª ed. Springer, Nueva York

Nota del editor Springer Nature permanece neutral con respecto a los reclamos jurisdiccionales en mapas publicados y afiliaciones institucionales.