



Fake News Detection Web App

Verificación de autenticidad de artículos periodísticos



Fake News Detection Web App

1

Presentación General

Fake News Detection Web App



Fake News Detection Web App

Desarrolladores: **Data Science Team**

Manuela Saez - Nataniel Martinez - Luis Vasquez - Leonardo Ferreira da Silva - Luciano Darriba

Cliente: **Portal de Noticias**

Fake News Detection Web App

Problema del cliente:

- Baja credibilidad por el gran porcentaje de noticias que son falsamente seleccionadas como Fake News

Requerimientos del cliente:

- Herramienta para verificar la autenticidad de las noticias previo a su publicación, con el objetivo de distribuir información fiable y preservar la confianza de sus lectores.

Fake News Detection Web App

Solución propuesta:

Aplicación web para detectar noticias falsas. La detección se realiza utilizando algoritmos de machine learning.



Fake News Detection Web App

Etapas principales:

1. Desarrollo del **algoritmo** que **predice** si un artículo de noticias es falso o no.
2. Construcción de la **Aplicación Web**: Fácil de usar. Utiliza el algoritmo para identificar noticias falsas de manera precisa y retorna un resultado.

Fake News Detection Web App

Métricas (utilizando Random Forest):

F1 score \approx 90%

Equilibrio entre evitar FP (precisión) y detección de todas las fake news (recall)

El modelo predice de forma correcta aproximadamente el 90 % de las noticias, lo que resulta en un valor aceptable para su implementación.

Fake News Detection Web App

Conclusiones:

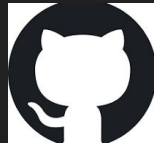
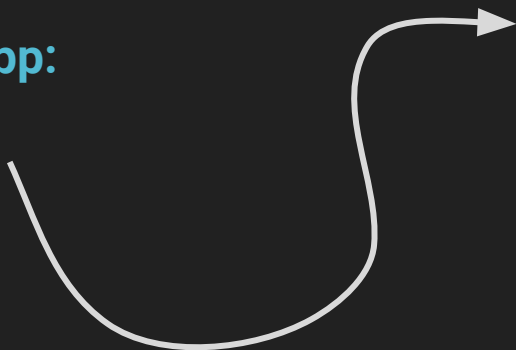
El modelo desarrollado cumple con los requerimientos del cliente y se encuentra en su etapa final de producción.

Integra: Técnicas avanzadas de aprendizaje automático + experiencia de usuario intuitiva.

Herramienta valiosa tanto para los **portales de noticias** como para los **lectores**. Promueve la diseminación de información precisa y confiable, ayudando a combatir la desinformación y mejorar la credibilidad del portal de comunicación.

Fake News Detection Web App

Uso de la WebApp:
Solo un click



[Descarga el repo aqui](#)

Fake News Text Classification

Selected Model: Random Forest

OPTION 1: Manual Data Entry

Author

Title

Text

Submit

OPTION 2: Upload a file

Upload CSV file



Drag and drop file here

Limit 200MB per file • CSV

Browse files

	author	news_count	percentage_real_news	status
0		0	0	New

Evaluate

Save Data

Fake News Detection Web App

2

Detalles Técnicos

Fake News Detection Web App

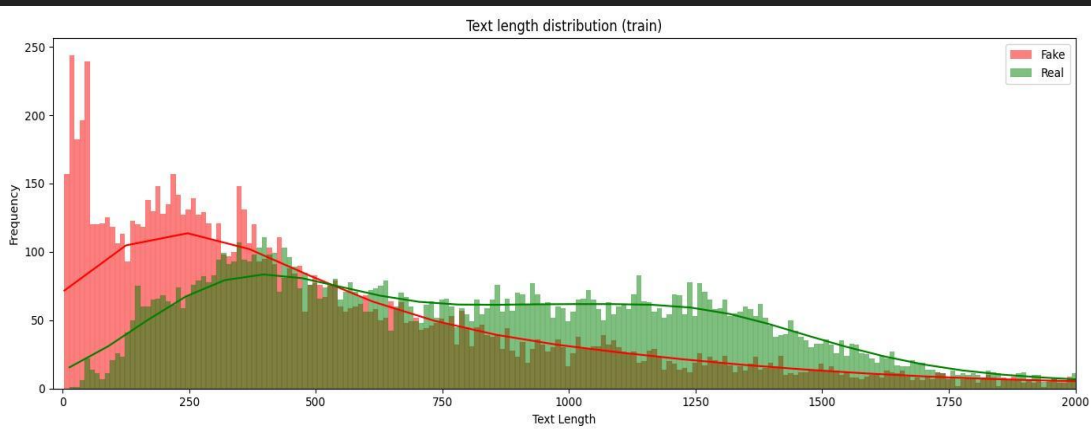
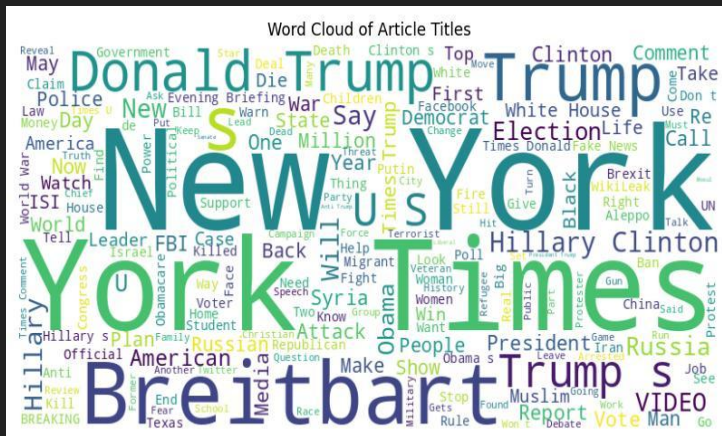
Detalles técnicos:



Fake News Detection Web App

Detalles técnicos: DATOS

- Train data set con ~21000 artículos etiquetados.
- Inglés
- Identificación de datos faltantes y limpieza

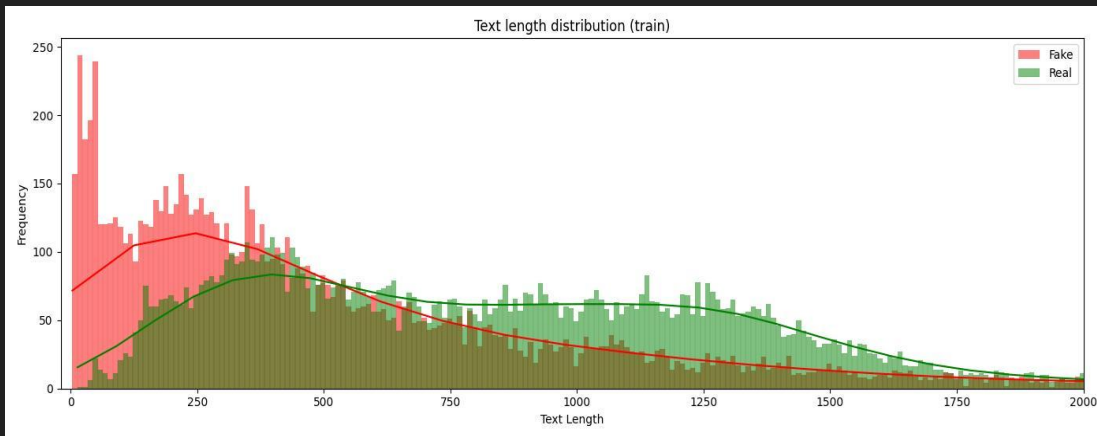
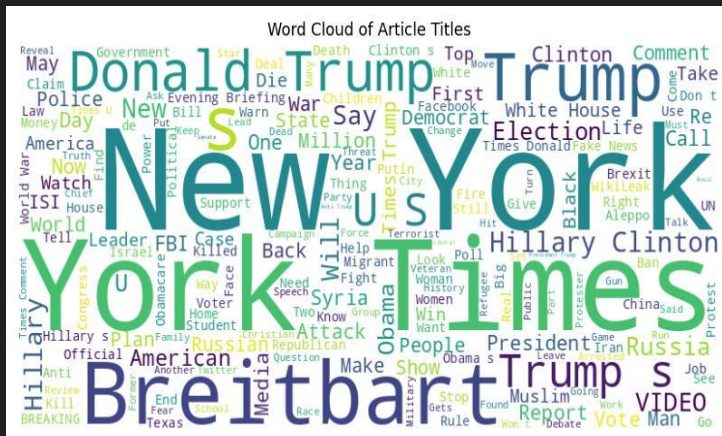


Fake News Detection Web App

Detalles técnicos: DATOS

- Train data set con ~21000 artículos etiquetados.
- Inglés
- Identificación de datos faltantes y limpieza

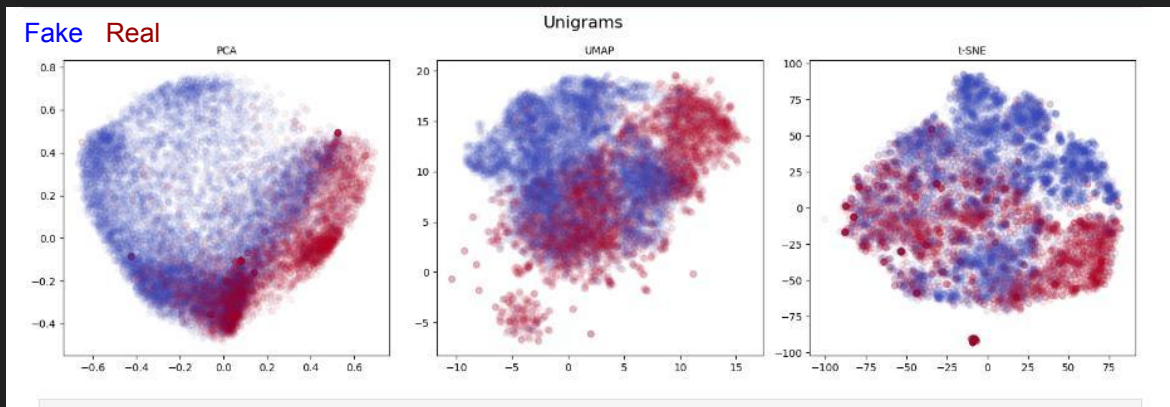
Pandas - NumPy - Langdetect -
Wordcloud - os - re



Fake News Detection Web App

Detalles técnicos: DATOS

- Procesamiento del texto: Tokenizacion + Lematizacion + Vectorizacion (uni, max_features)
- Feature extraction: train and validation dataset

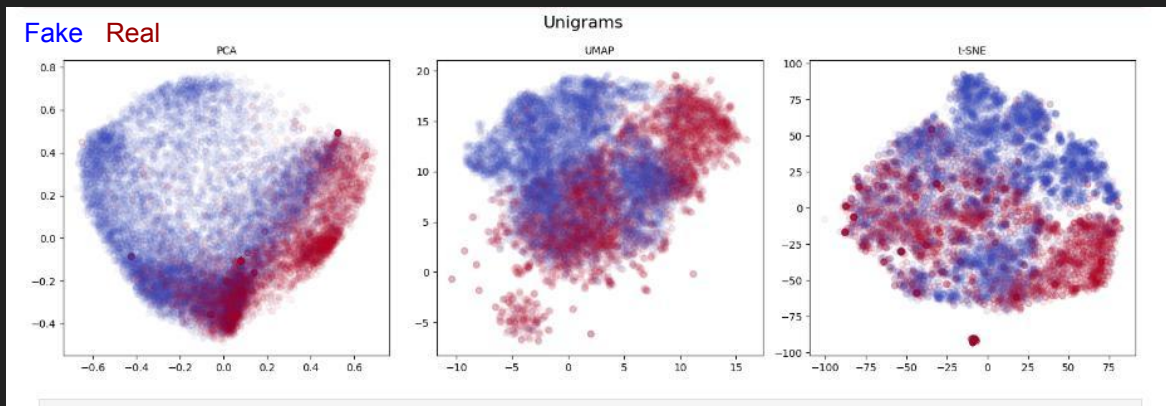


Fake News Detection Web App

Detalles técnicos: DATOS

- Procesamiento del texto: Tokenizacion + Lematizacion + Vectorizacion (uni, max_features)
- Feature extraction: train and validation dataset

Scikit-learn - TfidfVectorizer - Nltk - UMAP



Fake News Detection Web App

Detalles técnicos: ALGORITMO DE ML

- Supervisado, Clasificación
- 3 modelos: RandomForest, RegresionLogisitca, GaussianNaiveBayes
- Entrenamiento + performance + guardado de los modelos entrenados

RF

Accuracy: 0.915958873491283

Precision: 0.8962765957446809

Recall: 0.9030010718113612

F1 Score: 0.8996262680192205

RL

Accuracy: 0.8824318283415288

Precision: 0.8537486800422387

Recall: 0.8665594855305466

F1 Score: 0.8601063829787234

GNB

Accuracy: 0.8071077335717479

Precision: 0.8101295094701052

Recall: 0.8071077335717479

F1 Score: 0.8079030027325982

Fake News Detection Web App

Detalles técnicos: ALGORITMO DE ML

- Supervisado, Clasificación
- 3 modelos: RandomForest, RegresionLogisitca, GaussianNaiveBayes
- Entrenamiento + performance + guardado de los modelos entrenados

RF

Accuracy: 0.915958873491283

Precision: 0.8962765957446809

Recall: 0.9030010718113612

F1 Score: 0.8996262680192205

RL

Accuracy: 0.8824318283415288

Precision: 0.8537486800422387

Recall: 0.8665594855305466

F1 Score: 0.8601063829787234

GNB

Accuracy: 0.8071077335717479

Precision: 0.8101295094701052

Recall: 0.8071077335717479

F1 Score: 0.8079030027325982

Scikit-learn - Pickle

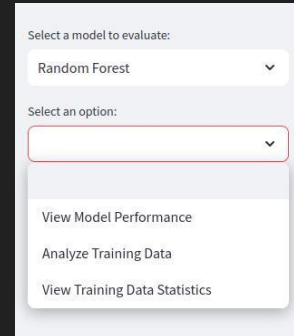
Fake News Detection Web App

Detalles técnicos: WEB APP

- Streamlit app
- Carga de datos: manual o CSV file (autor + título + artículo)
- 3 modelos de ML para seleccionar
- Una vez cargados los datos:
 - Muestra si hay datos previos del autor ingresado
 - Evalúa (predice) → etiqueta artículos como F o V.
 - Guarda los resultados → amplía futura database

Además:

- Performance distintos modelos
- Análisis de los datos de training
- Estadísticas de los datos de training
- Detección de noticias repetidas



The screenshot displays a web application interface with a light gray background. At the top, there is a label 'Select a model to evaluate:' followed by a dropdown menu currently showing 'Random Forest'. Below this is another label 'Select an option:' followed by an empty dropdown menu. At the bottom, there is a white rectangular box containing three links: 'View Model Performance', 'Analyze Training Data', and 'View Training Data Statistics'.

Fake News Detection Web App

Detalles técnicos: WEB APP

- Streamlit app
- Carga de datos: manual o CSV file (autor + título + artículo)
- 3 modelos de ML para seleccionar
- Una vez cargados los datos:
 - Muestra si hay datos previos del autor ingresado
 - Evalúa (predice) → etiqueta artículos como F o V.
 - Guarda los resultados → amplía futura database

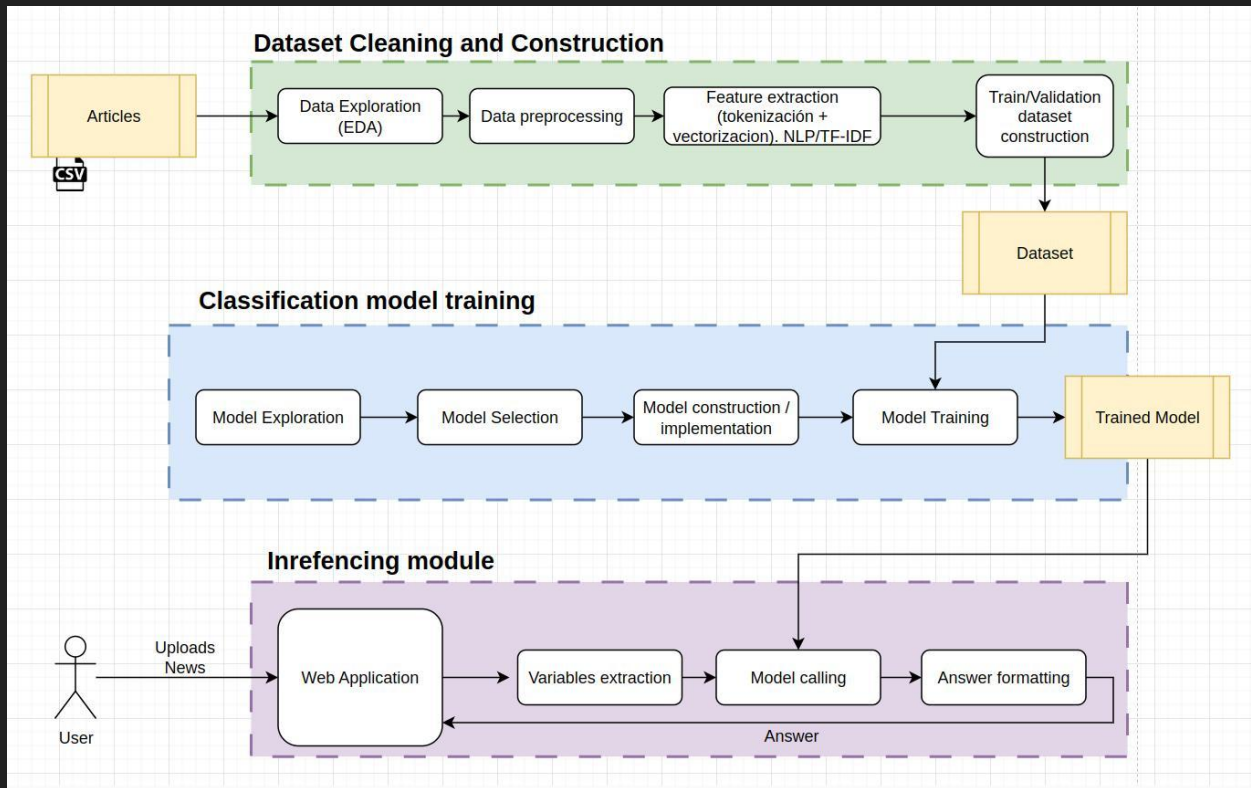
Además:

- Performance distintos modelos
- Análisis de los datos de training
- Estadísticas de los datos de training
- Detección de noticias repetidas

Scikit-learn - Pickle - Streamlite -
StreamlitCloud - GitHub

Fake News Detection Web App

Flujo técnico detallado de la solución:



Fake News Detection Web App

3

Demo de uso en vivo

Fake News Detection Web App

4

Perspectivas a corto y largo plazo

Fake News Detection Web App

Futuras mejoras a corto plazo:

- Ampliación del análisis estadístico por autor
- Profundizar en el análisis del vocabularios característico de las fake news
- Deploy en cloud
- Corregir Overfitting: regularización, cross-validation, reducción de dimensionalidad, aumento del tamaño de los datos, etc.
- Incorporación de datos adicionales
- Enriquecimiento del Dataset con nuevas características
- Experimentación con modelos más avanzados

Fake News Detection Web App

Futuras mejoras a largo plazo:

- Implementación en distintos idiomas
- Actualización periódicas utilizando el dataset de entrenamiento con las noticias recolectadas
- Segmentación particular para cada plataforma donde se suben las noticias
- Aplicación de forma local para ser utilizada por cada autor
- Mejoras en la Interfaz de Usuario

Fake News Detection Web App

Posibles Aplicaciones

1. Integración en Plataformas de Redes Sociales:

- **Descripción:** El modelo podría integrarse en plataformas de redes sociales para marcar automáticamente noticias potencialmente falsas antes de que se difundan.
- **Beneficio:** Ayuda a reducir la propagación de desinformación en redes sociales.

2. Extensión para Navegadores Web:

- **Descripción:** Desarrollar una extensión para navegadores que analice y etiquete las noticias mientras los usuarios navegan por la web.
- **Beneficio:** Proporciona a los usuarios una herramienta de verificación en tiempo real.

3. Asistente para Periodistas y Editores:

- **Descripción:** Implementar el modelo como una herramienta para periodistas y editores que les permita verificar la veracidad de la información antes de su publicación.
- **Beneficio:** Mejora la precisión y confiabilidad de las publicaciones periodísticas.

Fake News Detection Web App

Posibles Aplicaciones

4. Aplicación Educativa:

- **Descripción:** Crear una aplicación educativa que enseñe a los usuarios cómo identificar noticias falsas, usando ejemplos analizados por el modelo.
- **Beneficio:** Aumenta la alfabetización mediática y ayuda a los usuarios a desarrollar habilidades críticas para evaluar la información.

5. Monitoreo de Noticias en Tiempo Real:

- **Descripción:** Usar el modelo para monitorear noticias en tiempo real y alertar sobre posibles olas de desinformación durante eventos críticos (elecciones, crisis sanitarias, etc.).
- **Beneficio:** Proporciona una herramienta preventiva para evitar la propagación masiva de desinformación.

6. Servicio para Organizaciones y Gobiernos:

- **Descripción:** Ofrecer el modelo como un servicio a organizaciones gubernamentales o corporaciones para el monitoreo de la veracidad de la información en medios y redes sociales.
- **Beneficio:** Ayuda a estas organizaciones a tomar decisiones informadas basadas en datos confiables.