



ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS

INNOVACIÓN PARA LA EXCELENCIA

detectornoticiasfalsasecuador.com

Presentación de Proyecto de Tesis

Tema: Desarrollo de una herramienta web de detección de Fake News en Ecuador, empleando algoritmos de Machine Learning

Autor: Diakov Artem

Directora de tesis: Ing. Dorys Quiroz

INTRODUCCIÓN

FAKE NEWS

Auge de las redes sociales



Crecimiento de las fake news

Falsedad de la información



Daños a nivel social, político y económico

Ecuador no esta excepto



Lidiar con la propagación

Es esencial la investigación y desarrollo de una herramienta web de detección de Fake News en el país, utilizando algoritmos de Machine Learning

Objetivos

Objetivo General

Desarrollar una herramienta web de detección de Fake News en Ecuador utilizando Machine Learning, recopilando información de diferentes datasets libres

Objetivos específicos

- Obtener y recolectar información de Fake News y datasets libres.
- Transformar y traducir los datos obtenidos a idioma español.
- Diseñar el motor de detección de Fake News utilizando Inteligencia Artificial
- Desarrollar la aplicación web.
- Probar la herramienta con datos externos.

Hipótesis

La detección de Fake News en Ecuador utilizando Deep Learning o Machine Learning puede discernir información verdadera de la falsa

Marco Teórico

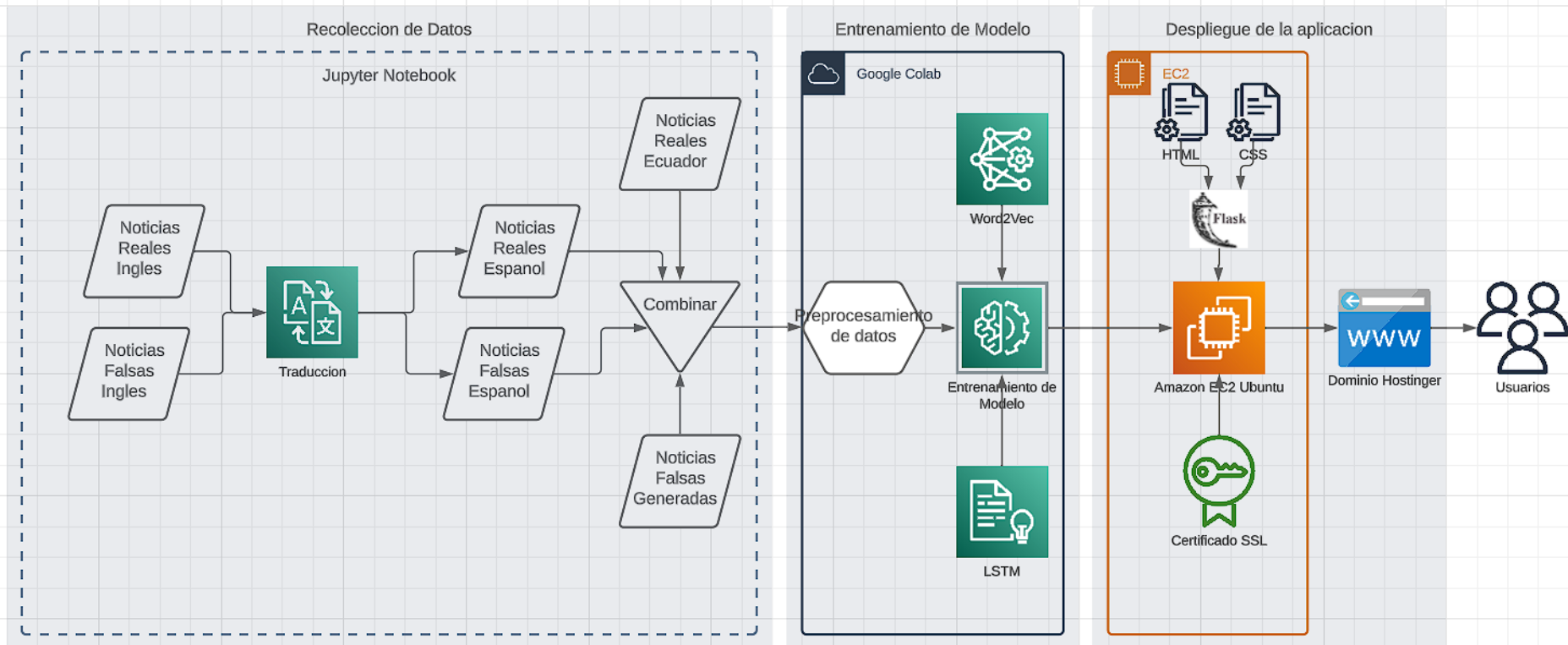
Fake News: Utilizadas para crear percepciones sobre la opinión pública

Herramientas Web: conjunto de aplicaciones, plataformas y tecnologías que facilitan el diseño, desarrollo, implementación y mantenimiento de sitios y aplicaciones web.

Machine Learning: subconjunto de Inteligencia Artificial que incluye técnicas estadísticas abstrusas que permiten a las máquinas mejorar en tareas

Desarrollo

Diagrama de modelo desarrollado



Traducción de noticias internacionales a español.



Hugging Face



Recolección de noticias de Ecuador



BeautifulSoup



Selenium WebDriver



Generación de noticias falsas usando AI



Noticias Reales
Traducidas

Noticias Falsas
Traducidas

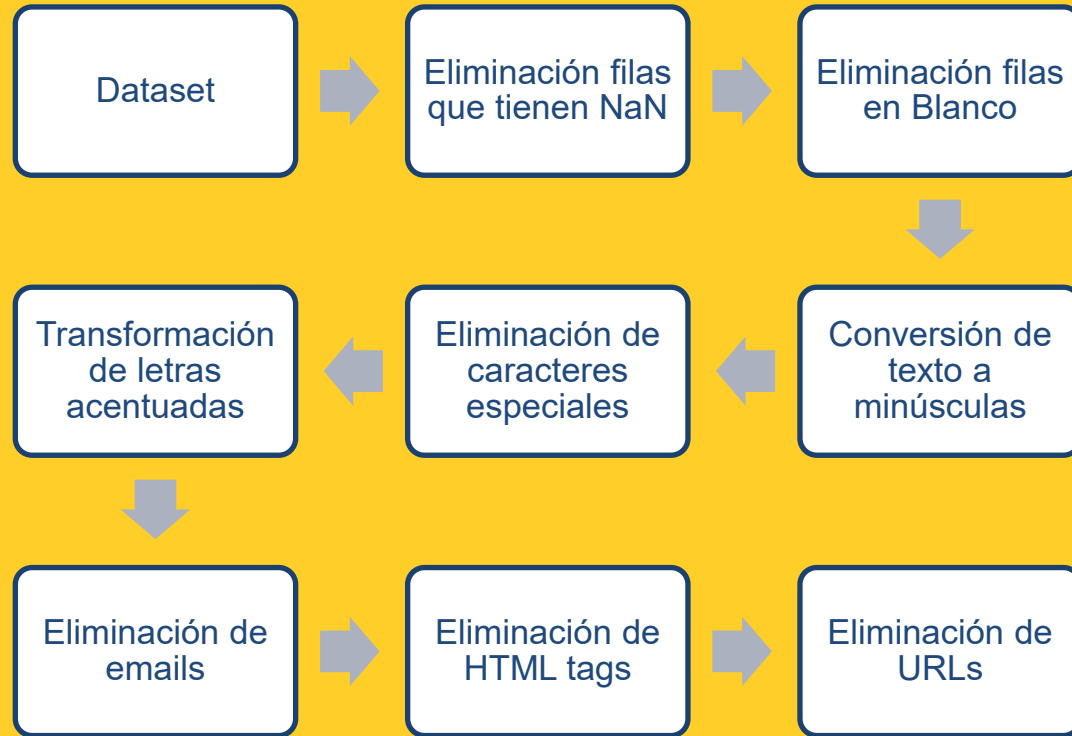
Dataset a
entrenar

Noticias
recolectadas de
Ecuador

Noticias
Generadas por
CHAT-GPT



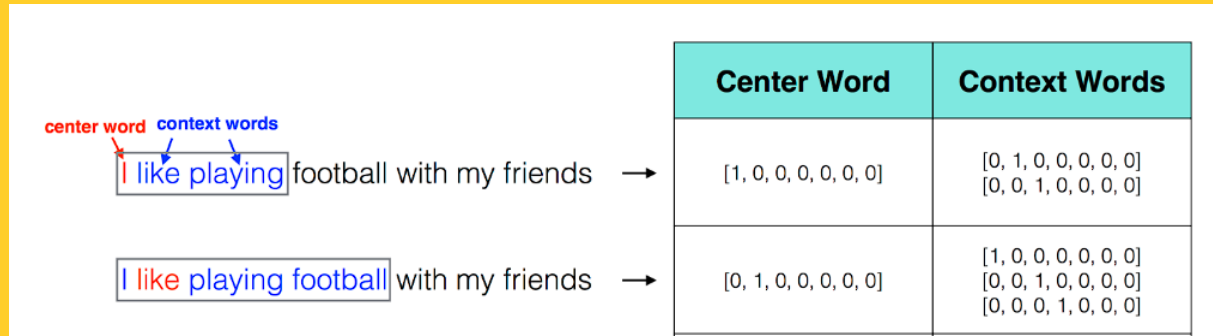
Preprocesamiento de datos.



Preprocesamiento de datos.



Vectorización de palabras: Word2Vec.



```
1 DIM = 100  
2 w2v_model = gensim.models.Word2Vec(sentences=X, vector_size=DIM, window=10, min_count=1)
```

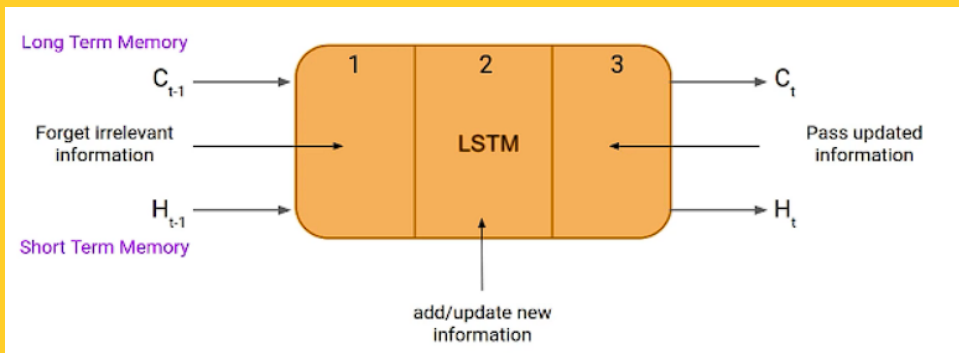
```
1 w2v_model.wv["ecuador"]  
  
array([ 2.9758513,  0.53403574, -3.212536,  0.2552427, -0.84223926,  
        1.6088291,  1.4793164, -1.016321, -0.33056805,  1.4050379,  
        1.3994269, -1.011899, -0.89514744, -0.30062824,  1.2992803,  
        -0.43870655,  2.827291, -0.71992296, -3.8615525,  0.485149,  
        -0.05570083, -0.50613654, -0.02754175,  0.03776176, -0.7459254,  
        0.3128388, -1.1514877,  0.09971294, -2.835829,  1.4448897,  
        1.3911756, -1.6043898,  0.06885928,  0.09047494,  2.131668,  
        0.2906076,  2.0475478, -0.10031062, -0.2563416,  1.1207207,  
        -2.1758544,  0.14934951, -0.2973839,  1.6523597, -1.7755029,  
        -1.9484416, -1.2201562,  0.22411971, -0.6127751,  2.692818, ]
```

```
2 w2v_model.wv.most_similar('correa')  
  
[(('glas', 0.8519338369369507),  
  ('leonidas', 0.8250226974487305),  
  ('iza', 0.8181368112564087),  
  ('glass', 0.812346875667572),  
  ('indigena', 0.6915594339370728),  
  ('corrupcioncorrea', 0.6382536292076111),  
  ('guillermo', 0.6368306279182434),  
  ('lasso', 0.6263988018035889),  
  ('poderademas', 0.6246264576911926),  
  ('rumorea', 0.6214133501052856)]
```



Entrenamiento de modelo usando LSTM

```
10 model = Sequential()  
11 model.add(Embedding(vocab_size, output_dim=DIM, weights = [embedding_vectors], input_length=maxlen, trainable=False))  
12 model.add(Bidirectional(LSTM(units=256)))
```



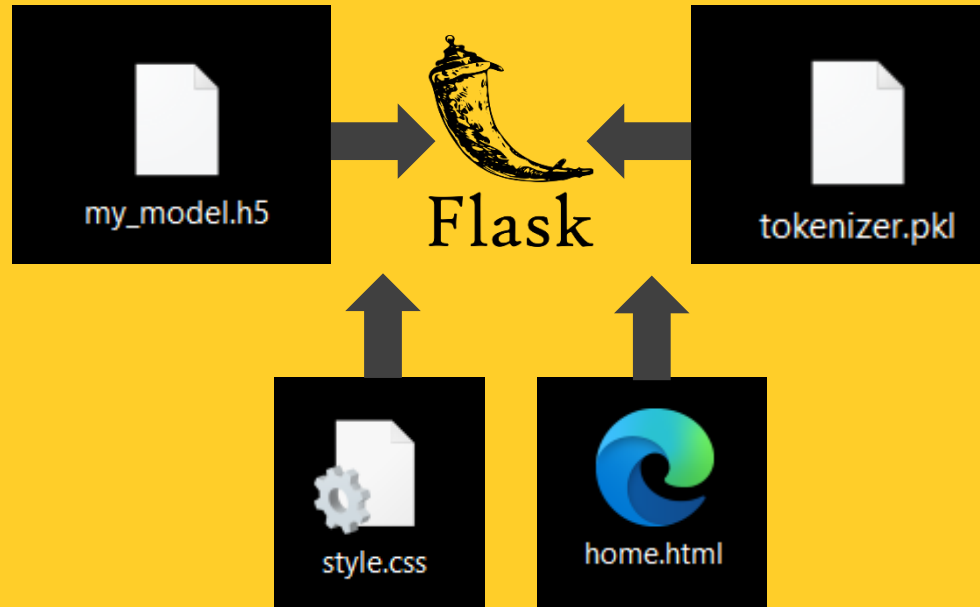
```
1 model.summary()
```

Model: "sequential"

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 1000, 1000)	422941000
bidirectional (Bidirectional)	(None, 512)	2574336
dense (Dense)	(None, 1)	513

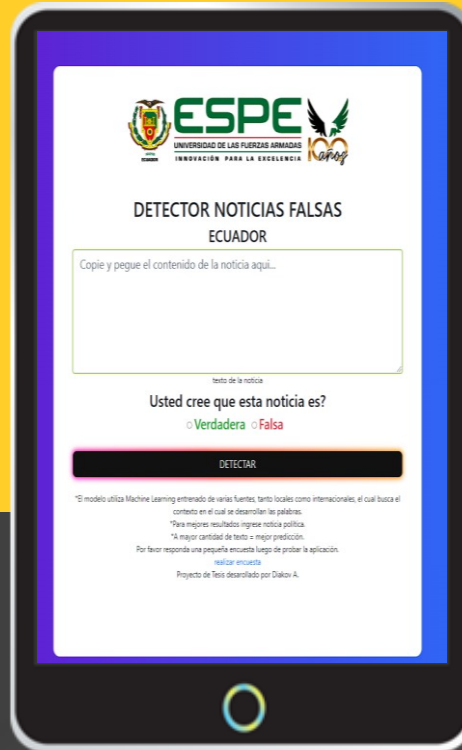
Total params: 425,515,849
Trainable params: 2,574,849
Non-trainable params: 422,941,000

Desarrollo de herramienta WEB

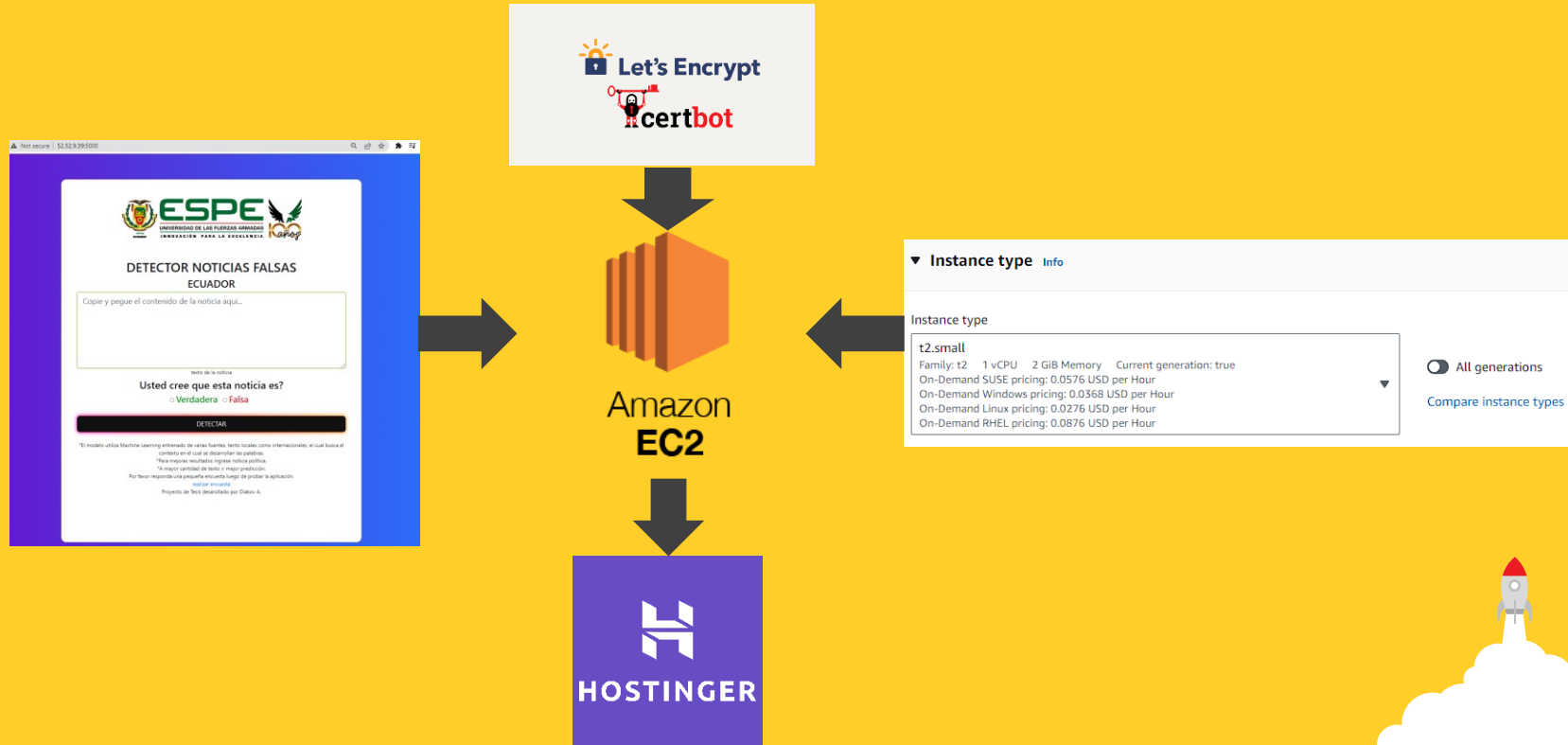


Desarrollo de herramienta WEB

Aplicación ejecutada



Despliegue de la aplicación

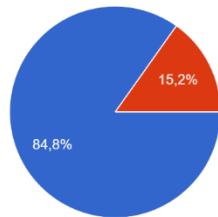


Encuestas realizadas a usuarios.

Respuesta media.

Pregunta #	Versión 1 (n=40)	Versión 2 (n=36)	Versión 3 (n=39)
1	4.17	4.05	4.15
2	4.10	3.66	4.03
3	3.80	3.64	3.64
4	4.03	3.91	3.92
5	3.87	3.94	3.84

UNICAMENTE RESPONDA SI HA PROBADO LOS ANTERIORES VERSIONES DE LA APLICACION. ¿A su percepción usted cree que hay una mejora notable respecto a las versiones anteriores de la aplicación?
33 respuestas



- SI, LA PREDICCIÓN ES MÁS PRECISA
- NO, LA PREDICCIÓN ES PEOR O IGUAL



Sugerencias por parte de usuarios.

¿Que recomendaciones, funcionalidades o mejoras desearía ver en la aplicación?

28 respuestas

En el ámbito grafico sería un nuevo tamaño o estilo para el texto de la pantalla

La verdad ninguna, todo esta perfecto

Muy buen identificador de noticias falsas

Mayor usabilidad

Menos caracteres en las descripción de las noticias

Ninguna

Falta entrenar el modelo..

La noticia usada en realidad era falsa y dio como verdadera con mas del 68%

Detallar mejor que tipo de texto se necesita...solo titulo de la noticia....titulo y contenido..... enlace de la noticia sería mas efectivo....

podría permitir ingresar el link

¿Que recomendaciones, funcionalidades o mejoras desearía ver en la aplicación?

28 respuestas

Me parece muy interesante, y por el momento no tengo ninguna recomendacion.

Creo que sería mas factible , pegar un vinculo directamente de donde sale la noticia , para evitar estar copiando la noticia.

La implementación de esta aplicación es interesante, es decir, prueba la veracidad de de una noticia, al momento de probar la aplicación con una noticia reciente, la aplicación me afirmó con un 70%. Puede ser que no esté altamente configurado en noticias recientes o que no tenga una recopilación de las noticias del mundo.

If topics are trending, it would be nice to see a "top 5" or "top 10" list of suspicious topics to help minimize the spread. Also If the source of the news (via URL) could be captured, further analysis could be conducted to rate the trustworthiness of so-called news sources.

Que funcione con enlaces

Debería detectar mejor si la noticia es verdadera además podría incorporar google para buscar noticias directamente



Conclusiones

- **Herramienta de Detección**
- **Recolección de Datos**
- **Traducción al Español**
- **Generación por medio de Chat-GPT Fake News**
- **Motor de Detección**
- **Interfaz de Usuario**
- **Pruebas con noticias locales de Ecuador**
- **Potencial Futuro**



Recomendaciones

- **Modelo de Predicción**
- **Calidad del Dataset**
- **Seguridad Web**
- **Interfaz de Usuario**
- **Preprocesamiento de Datos**
- **Retroalimentación y Capacitación**
- **Integración con Plataformas Sociales.**



Google

Buscar con Google Voy a tener suerte



Gracias por su atención!



detectornoticiasfalsasecuador.com

