Fake News Detection Dataset

Proyecto Final

Tania Michelle Rubí Rojas

Facultad de Ciencias, UNAM

8 de junio de 2020

Contenido

1. Objetivo
2. Materiales
3. Resultados
4. Conclusiones
5. Bibliografía

Objetivo

Dada una noticia, queremos saber si es real o falsa.
Para atacar este problema, utilizaremos dos *métodos* de clasificación diferentes:

1. Perceptrón Multicapa.
   Se explicará la arquitectura utilizada en esta red neuronal y su precisión al momento de clasificar noticias.
2. Algoritmo Pasivo-Agresivo.
   Se explicará cómo funciona y su precisión al momento de clasificar noticias.

Además, se discutirán brevemente las ventajas y desventajas que tienen uno sobre otro.

Conjunto de Datos

El conjunto de noticias que utlizaremos para entrenar nuestros modelos se obtuvo del sitio web *Kaggle*

`https://www.kaggle.com/c/fake-news/data`

Conjunto de Datos

el cual contiene dos archivos:

1. *train.csv*
   Contiene un conjunto de datos de entrenamiento con los siguientes atributos:
   - **id**: el *id* único de la noticia.
   - **title**: el título de la noticia.
   - **author**: el autor de la noticia.
   - **text**: el texto de la noticia (podría estar incompleto).
   - **label**: la etiqueta que clasifica la noticia como `REAL` (0) o `FAKE` (1).
2. *test.csv*
   Contiene un conjunto de datos con los mismos atributos que el archivo anterior, pero sin *label*.

Conjunto de Datos

	id	title	author	text	label
0	0	House Dem Aide: We Didn't Even See Comey's Let...	Darrell Lucus	House Dem Aide: We Didn't Even See Comey's Let...	1
1	1	FLYNN: Hillary Clinton, Big Woman on Campus - ...	Daniel J. Flynn	Ever get the feeling your life circles the rou...	0
2	2	Why the Truth Might Get You Fired	Consortiumnews.com	Why the Truth Might Get You Fired October 29, ...	1
3	3	15 Civilians Killed In Single US Airstrike Hav...	Jessica Purkiss	Videos 15 Civilians Killed In Single US Airstr...	1
4	4	Iranian woman jailed for fictional unpublished...	Howard Portnoy	Print \nAn Iranian woman has been sentenced to...	1
5	5	Jackie Mason: Hollywood Would Love Trump if He...	Daniel Nussbaum	In these trying times, Jackie Mason is the Voi...	0
6	6	Life: Life Of Luxury: Elton John's 6 Favorite ...	NaN	Ever wonder how Britain's most iconic pop pian...	1
7	7	Benoît Hamon Wins French Socialist Party's Pre...	Alissa J. Rubin	PARIS — France chose an idealistic, traditi...	0
8	8	Excerpts From a Draft Script for Donald Trump'...	NaN	Donald J. Trump is scheduled to make a highly ...	0
9	9	A Back-Channel Plan for Ukraine and Russia, Co...	Megan Twohey and Scott Shane	A week before Michael T. Flynn resigned as nat...	0

Preprocesamiento de datos

```
# Renombramos los datos de la variable 'label'.
df["label"] = df["label"].replace({0:'REAL', 1:'FAKE'})

# Obtenemos las etiquetas de cada uno de los articulos.
labels = df.label

# Visualizamos las primeras 10 etiquetas.
labels.head(10)
```

Preprocesamiento de Datos

Obtenemos los datos del atributo **label** modificados

```
0    FAKE
1    REAL
2    FAKE
3    FAKE
4    FAKE
5    REAL
6    FAKE
7    REAL
8    REAL
9    REAL
Name: label, dtype: object
```

Preprocesamiento de Datos

```
# Dividimos nuestro conjunto de datos con un split del 75-25
# para entrenamiento y prueba.
X_train, X_test, y_train, y_test =
    train_test_split(data_frame['text'],
                     labels,
                     test_size = 0.25,
                     random_state = 7)
```

Preprocesamiento de Datos

```
# Inicializamos un TfidfVectorizer.
v = TfidfVectorizer(stop_words = 'english',
                    max_df = 0.7)

# Entrenamos y transformamos el conjunto de entrenamiento.
v_train =
    v.fit_transform(X_train.apply(lambda x: np.str_(x)))

# Transformamos el conjunto de prueba.
v_test =
    v.transform(X_test.apply(lambda x: np.str_(x)))
```

Perceptrón Multicapa

Utilizamos un *MLP* para clasificar.

```
# Arquitectura.
mlpc = MLPClassifier(max_iter = 1000,
                     learning_rate_init = 0.0001,
                     hidden_layer_sizes = (2, 4))
```

Perceptrón Multicapa

```
# Entrenamos nuestra red.
mlpc_entrenamiento = mlpc.fit(v_train, y_train)

# Predecimos en el conjunto de prueba.
mlpc_prediction = mlpc.predict(v_test)

# Obtenemos la precision.
score_mlp = accuracy_score(y_test, mlpc_prediction)
print(f'Precision: {round(score_mlp * 100, 2)}%')
```

Precisión: 95.77%

Algoritmo Pasivo-Agresivo

Un algoritmo Pasivo-Agresivo funciona genéricamente con esta regla de actualización.

$$\begin{cases} \overline{w_{t+1}} = argmin_{\bar{w}} \frac{1}{2} \|\overline{w} - \overline{w_t}\|^2 + C\xi^2 \\ L(\bar{w}; x_t, y_t) \leq \xi \end{cases}$$

Algoritmo Pasivo-Agresivo

Después de resolver ambas condiciones de actualización, obtenemos la regla de actualización de forma cerrada:

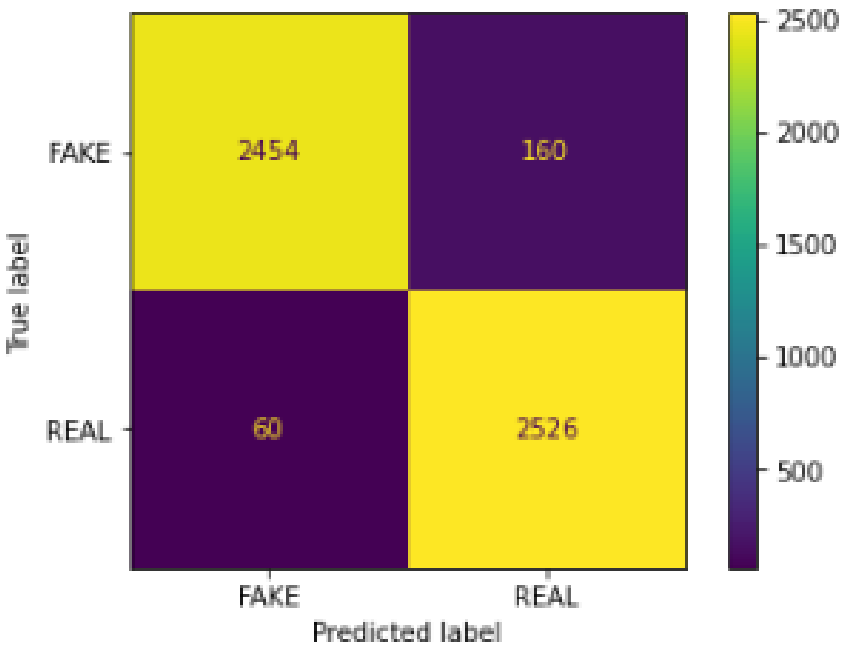
$$\overline{w_{t+1}} = \overline{w_t} + \frac{max(0, 1 - y_t(\overline{w^T} \cdot \overline{x_t}))}{\|x_t\|^2 + \frac{1}{2C}} y_t \overline{x_t}$$

Algoritmo Pasivo-Agresivo

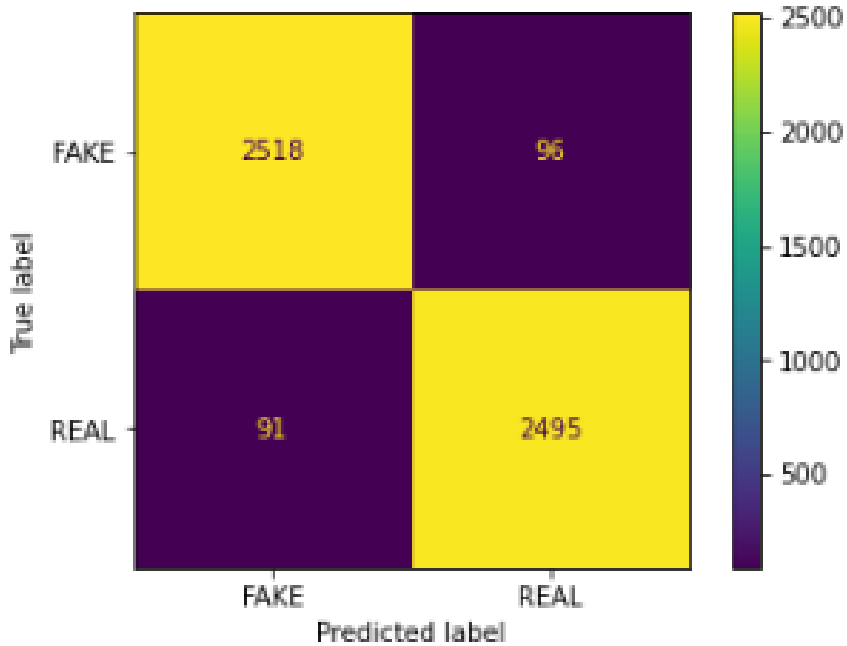
```
1  # Inicializamos un PassiveAggressiveClassifier.
2  pac = PassiveAggressiveClassifier(max_iter = 100)
3
4  # Entrenamos nuestro clasificador.
5  entrenamiento = pac.fit(v_train, y_train)
6
7  # Predecimos en el conjunto de prueba.
8  prediction_pac = pac.predict(v_test)
9
10 # Obtenemos la precision.
11 score = accuracy_score(y_test, prediction_pac)
12 print(f'Precision: {round(score * 100, 2)}%')
```

Precisión: 96.4%

Resultados



(a) MLP



(b) Algoritmo Pasivo-Agresivo

Figura: Matrices de confusión

Resultados

	precision	recall	f1-score	support
FAKE	0.98	0.94	0.96	2614
REAL	0.94	0.98	0.96	2586
accuracy			0.96	5200
macro avg	0.96	0.96	0.96	5200
weighted avg	0.96	0.96	0.96	5200

(a) MLP

	precision	recall	f1-score	support
FAKE	0.97	0.96	0.96	2614
REAL	0.96	0.96	0.96	2586
accuracy			0.96	5200
macro avg	0.96	0.96	0.96	5200
weighted avg	0.96	0.96	0.96	5200

(b) Algoritmo Pasivo-Agresivo

Figura: Reportes de clasificación

Resultados

	id	title	author	text
0	20800	Specter of Trump Loosens Tongues, if Not Purse...	David Streitfeld	PALO ALTO, Calif. — After years of scorning...
1	20801	Russian warships ready to strike terrorists ne...	NaN	Russian warships ready to strike terrorists ne...
2	20802	#NoDAPL: Native American Leaders Vow to Stay A...	Common Dreams	Videos #NoDAPL: Native American Leaders Vow to...
3	20803	Tim Tebow Will Attempt Another Comeback, This ...	Daniel Victor	If at first you don't succeed, try a different...
4	20804	Keiser Report: Meme Wars (E995)	Truth Broadcast Network	42 mins ago 1 Views 0 Comments 0 Likes 'For th...
5	20805	Trump is USA's antique hero. Clinton will be n...	NaN	Trump is USA's antique hero. Clinton will be n...
6	20806	Pelosi Calls for FBI Investigation to Find Out...	Pam Key	Sunday on NBC's "Meet the Press," House Minori...
7	20807	Weekly Featured Profile – Randy Shannon	Trevor Loudon	You are here: Home / *Articles of the Bound* /...
8	20808	Urban Population Booms Will Make Climate Chang...	NaN	Urban Population Booms Will Make Climate Chang...
9	20809	NaN	cognitive dissident	don't we have the receipt?

Resultados

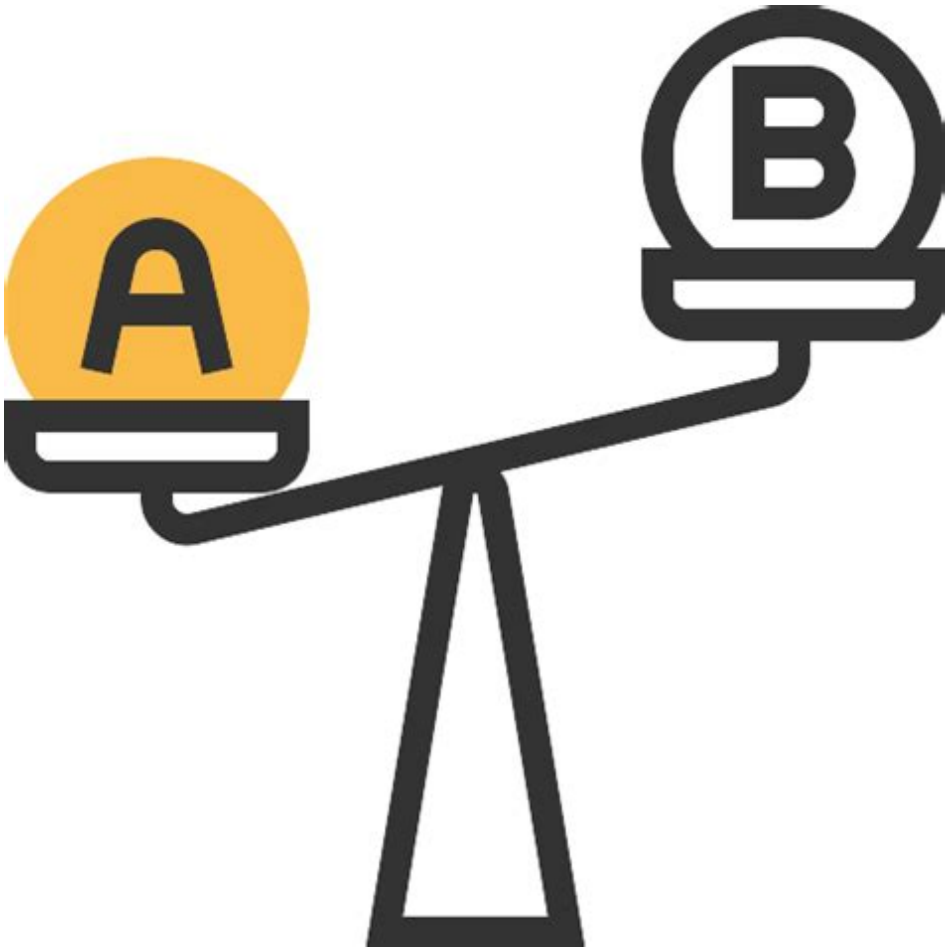
	title	label
0	Specter of Trump Loosens Tongues, if Not Purse...	FAKE
1	Russian warships ready to strike terrorists ne...	FAKE
2	#NoDAPL: Native American Leaders Vow to Stay A...	REAL
3	Tim Tebow Will Attempt Another Comeback, This ...	REAL
4	Keiser Report: Meme Wars (E995)	REAL
5	Trump is USA's antique hero. Clinton will be n...	REAL
6	Pelosi Calls for FBI Investigation to Find Out...	FAKE
7	Weekly Featured Profile – Randy Shannon	FAKE
8	Urban Population Booms Will Make Climate Chang...	REAL
9	NaN	FAKE
10	184 U.S. generals and admirals endorse Trump f...	REAL
11	"Working Class Hero" by John Brennon	REAL
12	The Rise of Mandatory Vaccinations Means the E...	REAL
13	Communists Terrorize Small Business	REAL
14	Computer Programmer Comes Forward, Admits To B...	REAL
15	Thieves Take a Chunk of Change, All 221 Pounds...	FAKE
16	New England Patriots' Owner, Still Sore at N.F...	FAKE

(a) MLP

	title	label
0	Specter of Trump Loosens Tongues, if Not Purse...	FAKE
1	Russian warships ready to strike terrorists ne...	FAKE
2	#NoDAPL: Native American Leaders Vow to Stay A...	REAL
3	Tim Tebow Will Attempt Another Comeback, This ...	REAL
4	Keiser Report: Meme Wars (E995)	REAL
5	Trump is USA's antique hero. Clinton will be n...	REAL
6	Pelosi Calls for FBI Investigation to Find Out...	FAKE
7	Weekly Featured Profile – Randy Shannon	FAKE
8	Urban Population Booms Will Make Climate Chang...	REAL
9	NaN	FAKE
10	184 U.S. generals and admirals endorse Trump f...	REAL
11	"Working Class Hero" by John Brennon	REAL
12	The Rise of Mandatory Vaccinations Means the E...	REAL
13	Communists Terrorize Small Business	REAL
14	Computer Programmer Comes Forward, Admits To B...	REAL
15	Thieves Take a Chunk of Change, All 221 Pounds...	FAKE
16	New England Patriots' Owner, Still Sore at N.F...	FAKE

(b) Algoritmo Pasivo-Agresivo

Conclusiones



Bibliografía

- BBC Noticias
  https://www.bbc.com/mundo/noticias-37910450.
  Redacción BBC
- ColombiaCheck
  https://colombiacheck.com/investigaciones/explicador-que-son-las-noticias-falsas
  Luisa Fernanda Gómez y José Manuel Cuevas
- Xataka
  https://www.xataka.com/otros/13-noticias-falsas-que-hemos-ayudado-a-difundir-por-interne

Bibliografía

- Kaggle: Fake News
  https://www.kaggle.com/c/fake-news/data
- Giuseppe Bonaccorso
  https://www.bonaccorso.eu/2017/10/06/ml-algorithms-addendum-passive-aggressive-algorithms/?subscribe=success#blog_subscription-2
- Stackbuse: Text classification with Python and Scikit Learn
  https://stackabuse.com/text-classification-with-python-and-scikit-learn/

Bibliografia

- Scikit Learn: Working with text data
  https://scikit-learn.org/stable/tutorial/text_analytics/working_with_text_data.html
- Kavita Ganesan: Tfidtransformer and TfidVectorizer: usage and differences
  https://kavita-ganesan.com/tfidftransformer-tfidfvectorizer-usage-differences/
- LUCA: Matriz de confusión
  https://empresas.blogthinkbig.com/ml-a-tu-alcance-matriz-confusion/