

The background of the slide is a light gray gradient. It is decorated with several realistic water droplets of various sizes. Some droplets are in the top left corner, others are scattered in the bottom right, and a few are near the center. The droplets have highlights and shadows, giving them a three-dimensional appearance.

FAKE NEWS **DETECTION**



+ Code + Text

✓ RAM
Disk

Editing

▼ Importing libraries

```
[1] import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.metrics import classification_report
import re
import string
```

```
[2] from google.colab import drive
drive.mount("/content/drive")
```

Mounted at /content/drive

▼ Importing Dataset

```
[4] path1 = "/content/drive/MyDrive/Fake.csv"
path2 = "/content/drive/MyDrive/True.csv"
```

+ Code + Text

✓ RAM  Disk  Editing ^

✓ [59] df_fake=pd.read_csv(path1)
df_true=pd.read_csv(path2)

✓ [6] df_fake.head()

	title	text	subject	date
0	Donald Trump Sends Out Embarrassing New Year'...	Donald Trump just couldn t wish all Americans ...	News	December 31, 2017
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017

✓ [7] df_true.head(5)

	title	text	subject	date
0	As U.S. budget fight looms, Republicans flip t...	WASHINGTON (Reuters) - The head of a conservat...	politicsNews	December 31, 2017
1	U.S. military to accept transgender recruits o...	WASHINGTON (Reuters) - Transgender people will...	politicsNews	December 29, 2017
2	Senior U.S. Republican senator: 'Let Mr. Muell...	WASHINGTON (Reuters) - The special counsel inv...	politicsNews	December 31, 2017
3	FBI Russia probe helped by Australian diplomat...	WASHINGTON (Reuters) - Trump campaign adviser ...	politicsNews	December 30, 2017
4	Trump wants Postal Service to charge 'much mor...	SEATTLE/WASHINGTON (Reuters) - President Donal...	politicsNews	December 29, 2017

+ Code + Text

✓ RAM  Disk  Editing 

▼ Importing a column "class" as target feature

```
✓ [9] df_fake["class"]=0  
0s df_true["class"]=1
```

```
✓ [10] df_fake.shape,df_true.shape  
0s ((23481, 5), (21417, 5))
```

```
✓ [11] # Removing last 10 rows for manual testing  
0s df_fake_manual_testing =df_fake.tail(10)  
for i in range (23480,23470,-1):  
    df_fake.drop([i],axis=0,inplace = True)  
  
df_true_manual_testing =df_true.tail(10)  
for i in range(21416,21406,-1):  
    df_true.drop([i],axis=0,inplace= True)
```

```
✓ [12] df_fake.shape,df_true.shape  
0s ((23471, 5), (21407, 5))
```

```
✓ [13] df_fake_manual_testing["class"]= 0  
0s df_true_manual_testing["class"]= 1
```

+ Code + Text

✓ RAM  Disk  Editing 

✓ [14] `df_fake_manual_testing.head(10)`

	title	text	subject	date	class
23471	Seven Iranians freed in the prisoner swap have...	21st Century Wire says This week, the historic...	Middle-east	January 20, 2016	0
23472	#Hashtag Hell & The Fake Left	By Dady Chery and Gilbert MercierAll writers ...	Middle-east	January 19, 2016	0
23473	Astroturfing: Journalist Reveals Brainwashing ...	Vic Bishop Waking TimesOur reality is carefull...	Middle-east	January 19, 2016	0
23474	The New American Century: An Era of Fraud	Paul Craig RobertsIn the last years of the 20t...	Middle-east	January 19, 2016	0
23475	Hillary Clinton: 'Israel First' (and no peace ...	Robert Fantina CounterpunchAlthough the United...	Middle-east	January 18, 2016	0
23476	McPain: John McCain Furious That Iran Treated ...	21st Century Wire says As 21WIRE reported earl...	Middle-east	January 16, 2016	0
23477	JUSTICE? Yahoo Settles E-mail Privacy Class-ac...	21st Century Wire says It s a familiar theme. ...	Middle-east	January 16, 2016	0
23478	Sunnistan: US and Allied 'Safe Zone' Plan to T...	Patrick Henningsen 21st Century WireRemember ...	Middle-east	January 15, 2016	0
23479	How to Blow \$700 Million: Al Jazeera America F...	21st Century Wire says Al Jazeera America will...	Middle-east	January 14, 2016	0
23480	10 U.S. Navy Sailors Held by Iranian Military ...	21st Century Wire says As 21WIRE predicted in ...	Middle-east	January 12, 2016	0

✓ [15] `df_true_manual_testing.head(10)`

	title	text	subject	date	class
21407	Mata Pires, owner of embattled Brazil builder ...	SAO PAULO (Reuters) - Cesar Mata Pires, the ow...	worldnews	August 22, 2017	1
21408	U.S., North Korea clash at U.N. forum over nuc...	GENEVA (Reuters) - North Korea and the United ...	worldnews	August 22, 2017	1
21409	U.S., North Korea clash at U.N. arms forum on ...	GENEVA (Reuters) - North Korea and the United ...	worldnews	August 22, 2017	1



✓ [15]	21415	Vatican upbeat on possibility of Pope Francis ...	MOSCOW (Reuters) - Vatican Secretary of State ...	worldnews	August 22, 2017	1
	21416	Indonesia to buy \$1.14 billion worth of Russia...	JAKARTA (Reuters) - Indonesia will buy 11 Sukh...	worldnews	August 22, 2017	1

{x}

▾ Merging True and Fake DataFrame

```
✓ [17] df_merge = pd.concat([df_fake,df_true],axis=0)
df_merge.head(10)
```

	title	text	subject	date	class	
0	Donald Trump Sends Out Embarrassing New Year'...	Donald Trump just couldn t wish all Americans ...	News	December 31, 2017	0	
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017	0	
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017	0	
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017	0	
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017	0	
5	Racist Alabama Cops Brutalize Black Boy While...	The number of cases of cops brutalizing and ki...	News	December 25, 2017	0	
6	Fresh Off The Golf Course, Trump Lashes Out A...	Donald Trump spent a good portion of his day a...	News	December 23, 2017	0	
7	Trump Said Some INSANELY Racist Stuff Inside ...	In the wake of yet another court decision that...	News	December 23, 2017	0	
8	Former CIA Director Slams Trump Over UN Bully...	Many people have raised the alarm regarding th...	News	December 22, 2017	0	
9	WATCH: Brand-New Pro-Trump Ad Features So Muc...	Just when you might have thought we d get a br...	News	December 21, 2017	0	

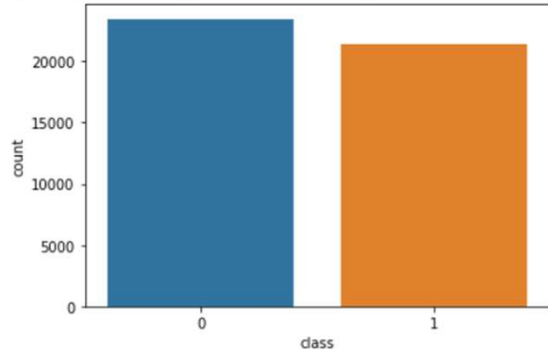


+ Code + Text

9 WATCH: Brand-new Pro-Trump Ad Features 30 M... Just when you might have thought we'd get a br... News December 21, 2017 0

✓ [61] sns.countplot(x="class", data=df_merge)

<matplotlib.axes._subplots.AxesSubplot at 0x7fcf6b2f1670>



✓ [18] df_merge.columns

Index(['title', 'text', 'subject', 'date', 'class'], dtype='object')

<> ▾ Removing column which are not required

✓ [20] df = df_merge.drop(["title", "subject", "date"], axis = 1)

+ Code + Text

✓ RAM Disk Editing ^

Removing column which are not required

```
[20] df = df_merge.drop(["title", "subject", "date"], axis = 1)
```

```
[21] df.isnull().sum()
```

```
text      0  
class     0  
dtype: int64
```

Random shuffling the DataFrame

```
[23] df = df.sample(frac = 1)
```

```
[24] df.head()
```

	text	class
11629	A Nashville, TN business fired 18 of its emplo...	0
17078	BOGOTA (Reuters) - Armed gunmen this week kill...	1
5389	(Reuters) - Highlights of the day for U.S. Pre...	1

✓ [24] 0s

11629	A Nashville, TN business fired 18 of its emplo...	0
17078	BOGOTA (Reuters) - Armed gunmen this week kill...	1
5389	(Reuters) - Highlights of the day for U.S. Pre...	1
14447	President Barack Obama and First Lady Michelle...	0
19528	Riots erupted in Charlotte after police office...	0

✓ [25] 0s

```
df.reset_index(inplace = True)
df.drop(["index"],axis=1,inplace=True)
```

✓ [26] 0s

```
df.columns
```

Index(['text', 'class'], dtype='object')

✓ [27] 0s

	text	class
0	A Nashville, TN business fired 18 of its emplo...	0
1	BOGOTA (Reuters) - Armed gunmen this week kill...	1
2	(Reuters) - Highlights of the day for U.S. Pre...	1
3	President Barack Obama and First Lady Michelle...	0
4	Riots erupted in Charlotte after police office...	0



+ Code + Text

✓ RAM
Disk

Editing



4 Riots erupted in Charlotte after police office... 0

▼ Creating a function to process the texts

```
[29] def wordopt(text):  
    text = text.lower()  
    text = re.sub('[.!?]', '', text)  
    text = re.sub("\W", " ", text)  
    text = re.sub('https?://\S+|www.\S+', '', text)  
    text = re.sub('<.+?>+', '', text)  
    text = re.sub('[%s]' % re.escape(string.punctuation), '', text)  
    text = re.sub('\n', '', text)  
    text = re.sub('\w*\d\w*', '', text)  
    return text
```

```
[30] df["text"] = df["text"].apply(wordopt)
```

▼ Defining dependent and independent variable

```
[32] x = df["text"]  
     y = df["class"]
```

4s completed at 10:10



+ Code + Text

✓ RAM  Disk  Editing ^

▼ Splitting Training and Testing

✓ [34] x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.25)
0s

▼ Convert text to vectors

✓ [36] from sklearn.feature_extraction.text import TfidfVectorizer
13s
vectorization = TfidfVectorizer()
xv_train = vectorization.fit_transform(x_train)
xv_test = vectorization.transform(x_test)

▼ Logistic Regression

✓ [38] from sklearn.linear_model import LogisticRegression
7s
LR = LogisticRegression()
LR.fit(xv_train, y_train)

<>
LogisticRegression()


✓ [39] pred_lr=LR.predict(xv_test)
0s

✓ 4s completed at 19:10



+ Code + Text

✓ RAM  Disk  Editing 

✓ 0s  LR.score(xv_test,y_test)

 0.9876114081996435

✓ 0s [41] `print(classification_report(y_test,pred_lr))`

	precision	recall	f1-score	support
0	0.99	0.99	0.99	5814
1	0.99	0.99	0.99	5406
accuracy			0.99	11220
macro avg	0.99	0.99	0.99	11220
weighted avg	0.99	0.99	0.99	11220

▼ Decision Tree Classification

✓ 37s [42] `from sklearn.tree import DecisionTreeClassifier`
`DT = DecisionTreeClassifier()`
`DT.fit(xv_train,y_train)`

`DecisionTreeClassifier()`

✓ 0s [43] `pred_dt = DT.predict(xv_test)`

✓ 4s completed at 10:10

+ Code + Text

✓ RAM  Disk  Editing 

✓ [44] DT.score(xv_test,y_test)
0.9950980392156863

✓ [45] print(classification_report(y_test,pred_dt))

	precision	recall	f1-score	support
0	0.99	1.00	1.00	5814
1	1.00	0.99	0.99	5406
accuracy			1.00	11220
macro avg	1.00	1.00	1.00	11220
weighted avg	1.00	1.00	1.00	11220

▾ Grading Boosting Classifier

✓ [47] from sklearn.ensemble import GradientBoostingClassifier

```
GBC = GradientBoostingClassifier(random_state=0)
GBC.fit(xv_train, y_train)
```

```
GradientBoostingClassifier(random_state=0)
```

✓ [48] pred_gbc = GBC.predict(xv_test)

✓ 4s completed at 19:10

+ Code + Text

✓ RAM  Disk  Editing 

✓ [49] GBC.score(xv_test,y_test)

0.9952762923351158

✓ [50] print(classification_report(y_test,pred_gbc))

	precision	recall	f1-score	support
0	1.00	0.99	1.00	5814
1	0.99	1.00	1.00	5406
accuracy			1.00	11220
macro avg	1.00	1.00	1.00	11220
weighted avg	1.00	1.00	1.00	11220

▼ Random Forest Classifier

✓ [52] from sklearn.ensemble import RandomForestClassifier

RFC = RandomForestClassifier(random_state=1)

RFC.fit(xv_train,y_train)

RandomForestClassifier(random_state=1)

✓ [53] pred rfc=RFC.predict(xv test)



```
randomForestClassifier(random_state=1)

[53] pred_rfc=RFC.predict(xv_test)

[54] RFC.score(xv_test,y_test)

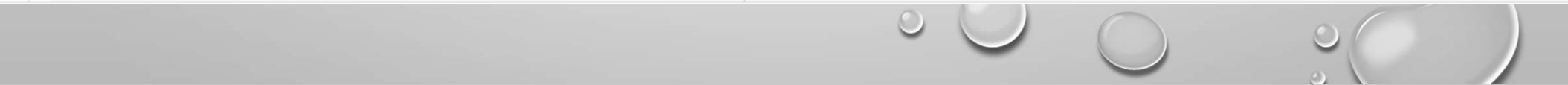
0.9906417112299465

[55] print(classification_report(y_test,pred_rfc))
```

	precision	recall	f1-score	support
0	0.99	0.99	0.99	5814
1	0.99	0.99	0.99	5406
accuracy			0.99	11220
macro avg	0.99	0.99	0.99	11220
weighted avg	0.99	0.99	0.99	11220

Model Testing

```
[57] def output_label(n):
    if n == 0:
        return "Fake News"
    elif n == 1:
        return "Not A Fake News"
```



+ Code + Text

✓ RAM  Disk  Editing ^

```
[57] return "Not A Fake News"

def manual_testing(news):
    testing_news = {"text":[news]}
    new_def_test = pd.DataFrame(testing_news)
    new_def_test["text"] = new_def_test["text"].apply(wordopt)
    new_x_test = new_def_test["text"]
    new_xv_test = vectorization.transform(new_x_test)
    pred_LR = LR.predict(new_xv_test)
    pred_DT = DT.predict(new_xv_test)
    pred_GBC = GBC.predict(new_xv_test)
    pred_RFC = RFC.predict(new_xv_test)

    return print("\n\nLR Prediction: {} \nDT Prediction: {} \nGBC Prediction: {} \nRFC Prediction: {}".format(output_label(pred_LR[0]),
                                                                 output_label(pred_GBC[0]),
                                                                 output_label(pred_RFC[0])))
```

```
[62] news = str(input())
      manual_testing(news)
```

Donald Trump Sends Out Embarrassing New Year's Eve Message; This is Disturbing

LR Prediction: Fake News
DT Prediction: Fake News
GBC Prediction: Fake News
RFC Prediction: Fake News

✓ 4s completed at 19:10