Date	23 October 2023
Team ID	NM2023TMID335
Team Name	Proj_227279_Team_1
Project Name	Building a smarter Ai-powered spam classifier

Building a Smarter AI-Powered Spam Classifier

Phase 4: Development Part 2

In this part you will continue building your project.

In this phase, we'll continue building our spam classifier by:

- Selecting a machine learning algorithm
- Training the model
- Evaluating its performance.

[1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score

import os
print(os.listdir(r"C:\Users\NELSON\Desktop\new deploy\vishwa"))
```

```
['.ipynb_checkpoints', 'Datst.csv', 'final', 'final.xlsx', 'M3.ipynb',
'M4.ipynb', 'M5.ipynb', 'M6.ipynb', 'prep.ipynb', 'preprocessing.ipynb',
'sanjay1.csv', 'sanjay2.csv', 'spam (4).csv', 'spamclassifier.csv',
'Untitled.ipynb']
```

[2]:
```
df = pd.read_csv(r"C:\Users\NELSON\Desktop\new deploy\vishwa/spamclassifier.
↪csv", encoding="latin-1")
```

[3]:
```
encoder = LabelEncoder()
df["v2_encode"] = encoder.fit_transform(df["v2"])
```

[4]:
```
X = df[["v2_encode"]]
y = df["v1"]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
↪random_state=42)
```

[5]:
```
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score
```

[6]:
```
model = RandomForestClassifier()
model.fit(X_train, y_train)
```

[6]: RandomForestClassifier()

[7]:
```
y_pred = model.predict(X_test)
score = accuracy_score(y_test, y_pred)
print("Model accuracy: ", score)
```

```
Model accuracy:  0.9040358744394619
```