```
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score
from sklearn.neighbors import KNeighborsClassifier


df = pd.read_csv("./emails.csv")


df.head()


df.isnull().sum()


X = df.iloc[:,1:3001]
X


Y = df.iloc[:,-1].values
Y
```

```
array([0, 0, 0, ..., 1, 1, 0], dtype=int64)
```

```
train_x,test_x,train_y,test_y = train_test_split(X,Y,test_size = 0.25)


svc = SVC(C=1.0,kernel='rbf',gamma='auto')
svc.fit(train_x,train_y)
y_pred2 = svc.predict(test_x)
print("Accuracy Score for SVC : ", accuracy_score(y_pred2,test_y))
```

```
Accuracy Score for SVC :  0.9033255993812839
```

```
X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size = 0.2, random_state=42)


knn = KNeighborsClassifier(n_neighbors=7)


knn.fit(X_train, y_train)
```

```
KNeighborsClassifier
KNeighborsClassifier(n_neighbors=7)
```

```
print(knn.predict(X_test))
```

```
[0 0 1 ... 0 1 0]
```

+ Code + Text

```
print(knn.score(X_test, y_test))
```

```
0.8685990338164251
```