

```
In [30]: import pandas as pd
```

```
In [31]: import numpy as np
```

```
In [52]: import matplotlib.pyplot as plt
```

```
In [33]: from sklearn.model_selection import GridSearchCV  
from sklearn import svm
```

```
In [34]: df = pd.read_csv(r'C:\Users\Rutu\Documents\New folder\spam.csv')
```

```
In [35]: df.head()
```

```
Out[35]:
```

	Label	EmailText
0	ham	Go until jurong point, crazy.. Available only ...
1	ham	Ok lar... Joking wif u oni...
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...
3	ham	U dun say so early hor... U c already then say...
4	ham	Nah I don't think he goes to usf, he lives aro...

```
In [36]: df.tail()
```

```
Out[36]:
```

	Label	EmailText
5567	spam	This is the 2nd time we have tried 2 contact u...
5568	ham	Will Æ b going to esplanade fr home?
5569	ham	Pity, * was in mood for that. So...any other s...
5570	ham	The guy did some bitching but I acted like i'd...
5571	ham	Rofl. Its true to its name

```
In [37]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 5572 entries, 0 to 5571  
Data columns (total 2 columns):  
#   Column      Non-Null Count  Dtype  
---  -  
0   Label       5572 non-null   object  
1   EmailText   5572 non-null   object  
dtypes: object(2)  
memory usage: 87.2+ KB
```

```
In [38]: df.isnull().sum()
```

```
Out[38]:
```

Label	0
EmailText	0
dtype:	int64

```
In [39]: df.dropna()
```

```
Out[39]:
```

	Label	EmailText
0	ham	Go until jurong point, crazy.. Available only ...

1	ham	Ok lar... Joking wif u oni...
2	spam	Free entry in 2 a wkly comp to win FA Cup fina...
3	ham	U dun say so early hor... U c already then say...
4	ham	Nah I don't think he goes to usf, he lives aro...
...	...	...
5567	spam	This is the 2nd time we have tried 2 contact u...
5568	ham	Will Å□_ b going to esplanade fr home?
5569	ham	Pity, * was in mood for that. So...any other s...
5570	ham	The guy did some bitching but I acted like i'd...
5571	ham	Rofl. Its true to its name

5572 rows × 2 columns

```
In [40]: df.describe()
```

```
Out[40]:
```

	Label	EmailText
count	5572	5572
unique	2	5169
top	ham	Sorry, I'll call later
freq	4825	30

```
In [41]: df.shape
```

```
Out[41]: (5572, 2)
```

```
In [ ]:
```

```
In [44]: X = df['EmailText']
y = df['Label']
```

```
In [45]: from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)
```

```
In [46]: from sklearn.feature_extraction.text import CountVectorizer
cv = CountVectorizer()
X_train = cv.fit_transform(X_train)
X_test = cv.transform(X_test)
```

```
In [47]: from sklearn.svm import SVC
classifier = SVC(kernel = 'rbf', random_state = 0)
classifier.fit(X_train, y_train)
```

```
Out[47]: SVC(random_state=0)
```

```
In [48]: print(classifier.score(X_test,y_test))
```

```
0.9766816143497757
```

```
In [ ]:
```

