C:\ProgramData\Anaconda3\envs\my\python.exe E:/\_USC/my/nlp-linkedin/nlp2.py

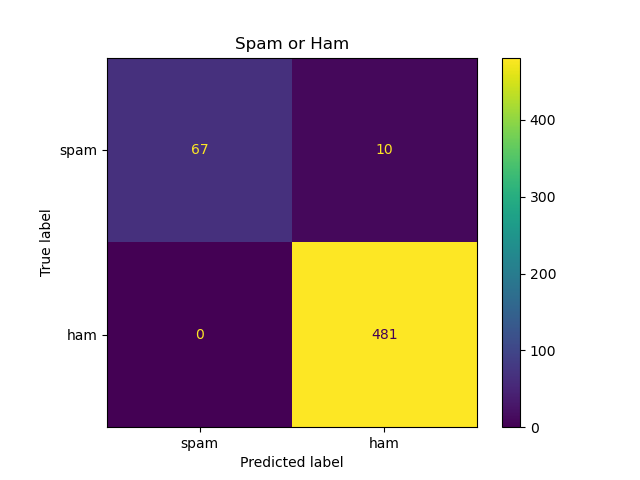
test acc= 0.982078853046595

bal test acc= 0.9898167006109979

F1 score= 0.982078853046595

F1 score= [0.93055556 0.98971193]

Process finished with exit code 0



C:\ProgramData\Anaconda3\envs\my\python.exe E:/\_USC/my/nlp-linkedin/nlp1.py

label body

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...

Dataset has 5572 rows and 2 col

747 spam and 4825 ham

nulls in label are 0

nulls in body are 0

label body len

0 ham Go until jurong point, crazy.. Available only ... 111

1 ham Ok lar... Joking wif u oni... 29

2 spam Free entry in 2 a wkly comp to win FA Cup fina... 155

3 ham U dun say so early hor... U c already then say... 49

4 ham Nah I don't think he goes to usf, he lives aro... 61

E:/\_USC/my/nlp-linkedin/nlp1.py:44: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

df['clean\_msg'][i]=message[i]

C:\ProgramData\Anaconda3\envs\my\lib\site-packages\pandas\core\indexing.py:1637: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

self.\_setitem\_single\_block(indexer, value, name)

label ... clean\_msg\_len

0 ham ... 76

1 ham ... 21

2 spam ... 99

3 ham ... 35

4 ham ... 36

[5 rows x 5 columns]

body len ... clean\_msg\_len label

0 Go until jurong point, crazy.. Available only ... 111 ... 76 1

1 Ok lar... Joking wif u oni... 29 ... 21 1

2 Free entry in 2 a wkly comp to win FA Cup fina... 155 ... 99 0

3 U dun say so early hor... U c already then say... 49 ... 35 1

4 Nah I don't think he goes to usf, he lives aro... 61 ... 36 1

[5 rows x 5 columns]

X\_train shape: (5014, 2500)

X\_test shape: (558, 2500)

Y\_train shape: (5014,)

Y\_test shape: (558,)

SVM's Accuracy:0.982078853046595

Process finished with exit code 0