

## ASSIGNMENT 4

Team ID : PNT2022TMID22506

DATE : 31 October, 2022

### Import the necessary libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from keras.models import Model
from keras.layers import LSTM, Activation, Dense, Dropout, Input, Embedding
from keras.optimizers import RMSprop
from keras.preprocessing.text import Tokenizer
from keras.preprocessing import sequence
from keras.utils import to_categorical
from keras.callbacks import EarlyStopping
%matplotlib inline
```

### Load the data into Pandas

```
df = pd.read_csv('/content/spam.csv', delimiter=',', encoding='latin-1')
df.head()
```

	v1	v2	Unnamed: 2	Unnamed: 3	Unnamed: 4
0	ham Go until jurong point, crazy.. Available only ...	NaN	NaN	NaN	
1	ham Ok lar... Joking wif u oni...	NaN	NaN	NaN	
2	spam Free entry in 2 a wkly comp to win FA Cup fina...	NaN	NaN	NaN	
3	ham U dun say so early hor... U c already then say...	NaN	NaN	NaN	
4	ham Nah I don't think he goes to usf, he lives aro...	NaN	NaN	NaN	

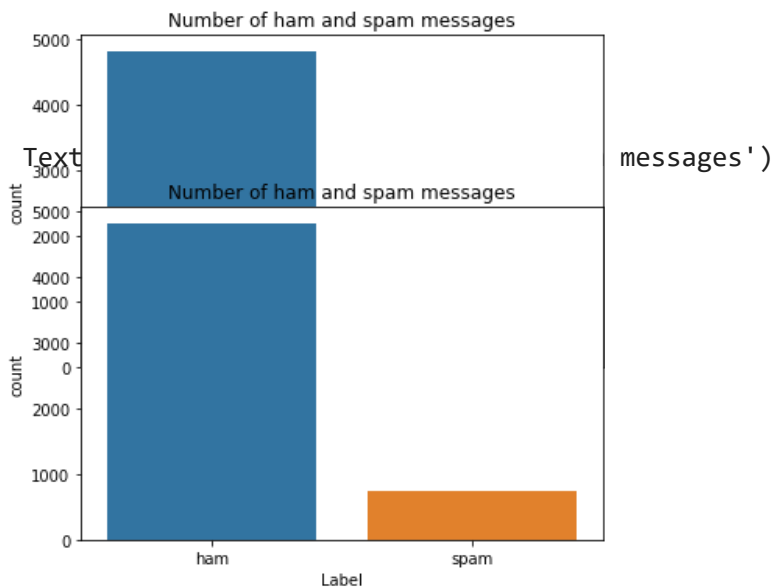
df.tail()

	v1	v2	Unnamed: 2	Unnamed: 3	Unnamed: 4
5567	spam This is the 2nd time we have tried 2 contact u...	NaN	NaN	NaN	
5568	ham Will Ì_ b going to esplanade fr home?	NaN	NaN	NaN	
5569	ham Pity, * was in mood for that. So...any other s...	NaN	NaN	NaN	
5570	ham The guy did some bitching but I acted like i'd...	NaN	NaN	NaN	
5571	ham Rofl. Its true to its name	NaN	NaN	NaN	

```

# Checking datatype df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5572 entries, 0 to 5571
Data columns (total 5 columns):
  # Column          Non-Null Count  Dtype
-----  -
0         v1         5572 non-null   object
1         v2         5572 non-null   object
2         Unnamed: 2     50 non-null    object
3         Unnamed: 3     12 non-null    object
4         Unnamed: 4     6 non-null     object
dtypes: object(5) memory usage: 217.8+ KB
df.drop(['Unnamed: 2', 'Unnamed: 3', 'Unnamed: 4'],axis=1,inplace=True)
df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5572 entries, 0 to 5571
Data columns (total 2 columns):
  # Column Non-Null Count  Dtype
-----  -
0         v1         5572 non-null   object
1         v2         5572 non-null   object
dtypes: object(2) memory usage: 87.2+ KB
sns.countplot(df.v1) plt.xlabel('Label')
plt.title('Number of ham and spam
messages')

```



```

/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43:
FutureWarning: Pass t FutureWarning
Text(0.5, 1.0, 'Number of ham and spam messages')

```

**# Convert list into array**

```
x_train,y_train = np.array(x_train),np.array(y_train)
```

**# Building model**

```
from tensorflow.keras.models import Sequential from
tensorflow.keras.layers import LSTM, Dense model = Sequential()
model.add(LSTM(50, input_shape=(60, 1),return_sequences=True))
model.add(LSTM(50,return_sequences=True))
model.add(LSTM(50,return_sequences=True))
model.add(LSTM(50,return_sequences=True)) model.add(Dense(1))
model.compile(optimizer='adam',loss='mse')
```

**# save model**

```
model.save('LSTM.h5')
```

