Spam Classification using Recurrent Neural Networks

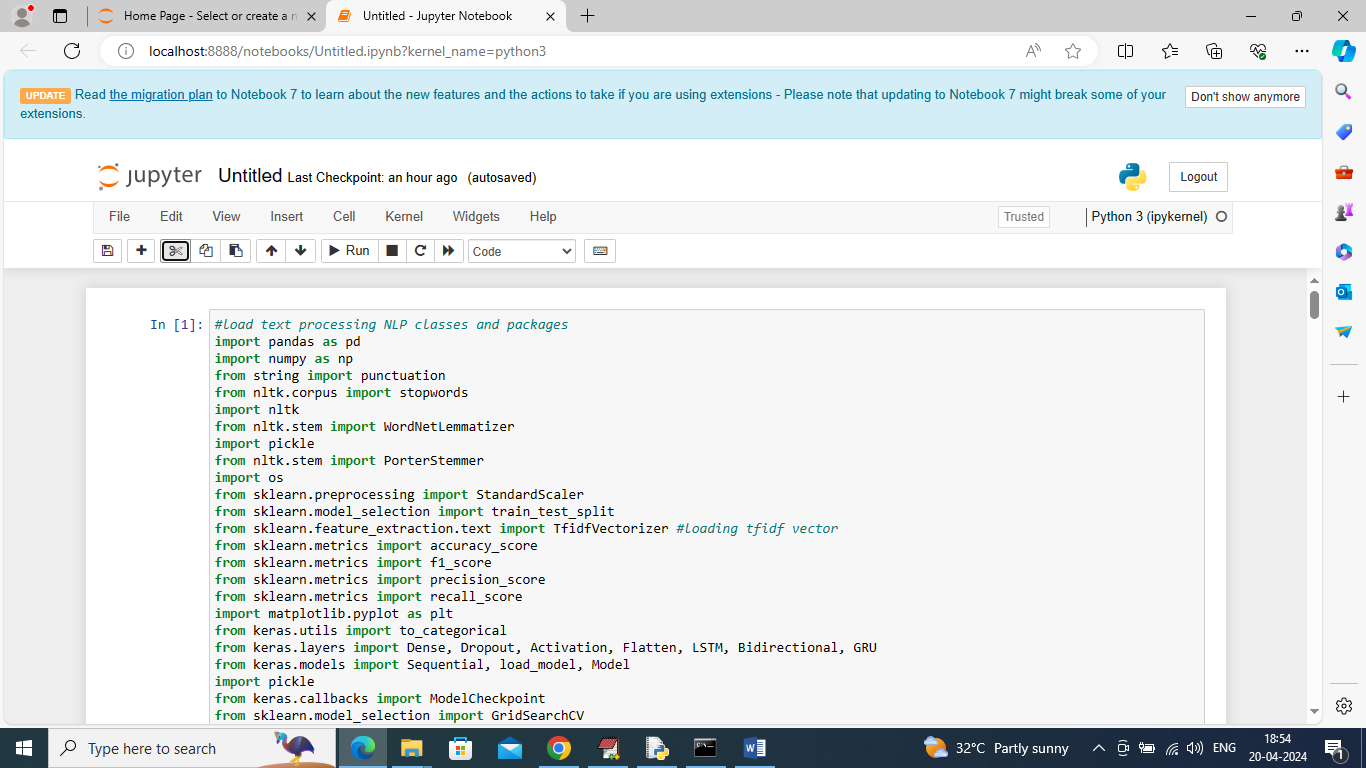
In today’s era growing technologies make human life very much simpler where all task can be perform virtually such as online shopping, intimating friends with short messages about meeting or other events but some malicious users taking advantages of this technologies to spread fake or spam messages. This spam messages can spread negativity in users which can lead them to avoid or stop usage of such services.

To detect such spam messages we are employing deep learning RNN recurrent neural network) algorithm which will get trained on past HAM and SPAM messages and this trained model can be applied on future coming messages to predict weather coming message is HAM or spam, if spam message detected then application will automatically take action and drop such messages.

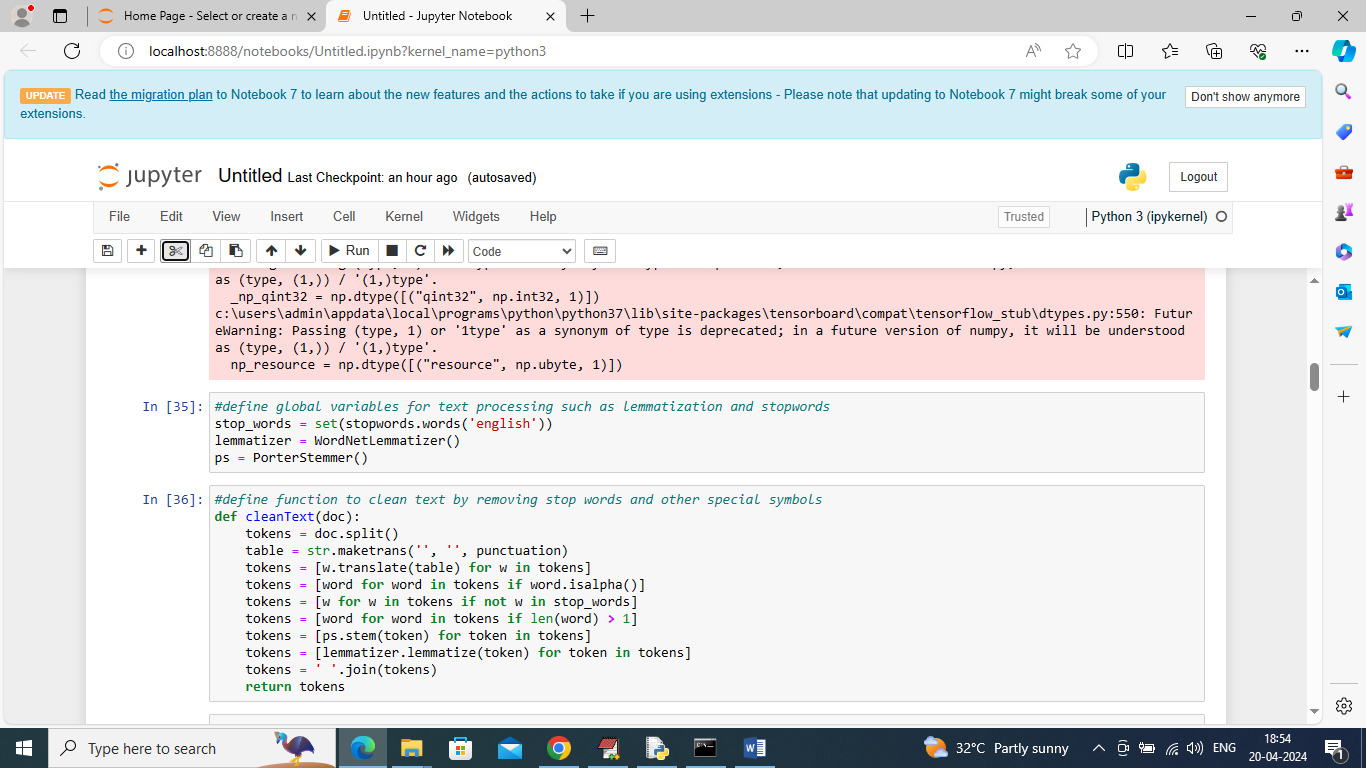
In propose work we have utilized LSTM (long short term memory) deep learning algorithm from RNN family to detect spam messages. LSTM widely used in Natural language and speech processing. LSTMs are able to process and analyse sequential data, such as time series, text, and speech. They use a memory cell and gates to control the flow of information, allowing them to selectively retain or discard information.

To further enhance LSTM performance we have employed LSTM tuning by using various hyper parameters such as training epochs and batch size and can increase many other parameters but training time will go longer so we have restrict ourselves with epochs and batch size. After tuning LSTM manages to classify spam messages with an accuracy of over 95%.

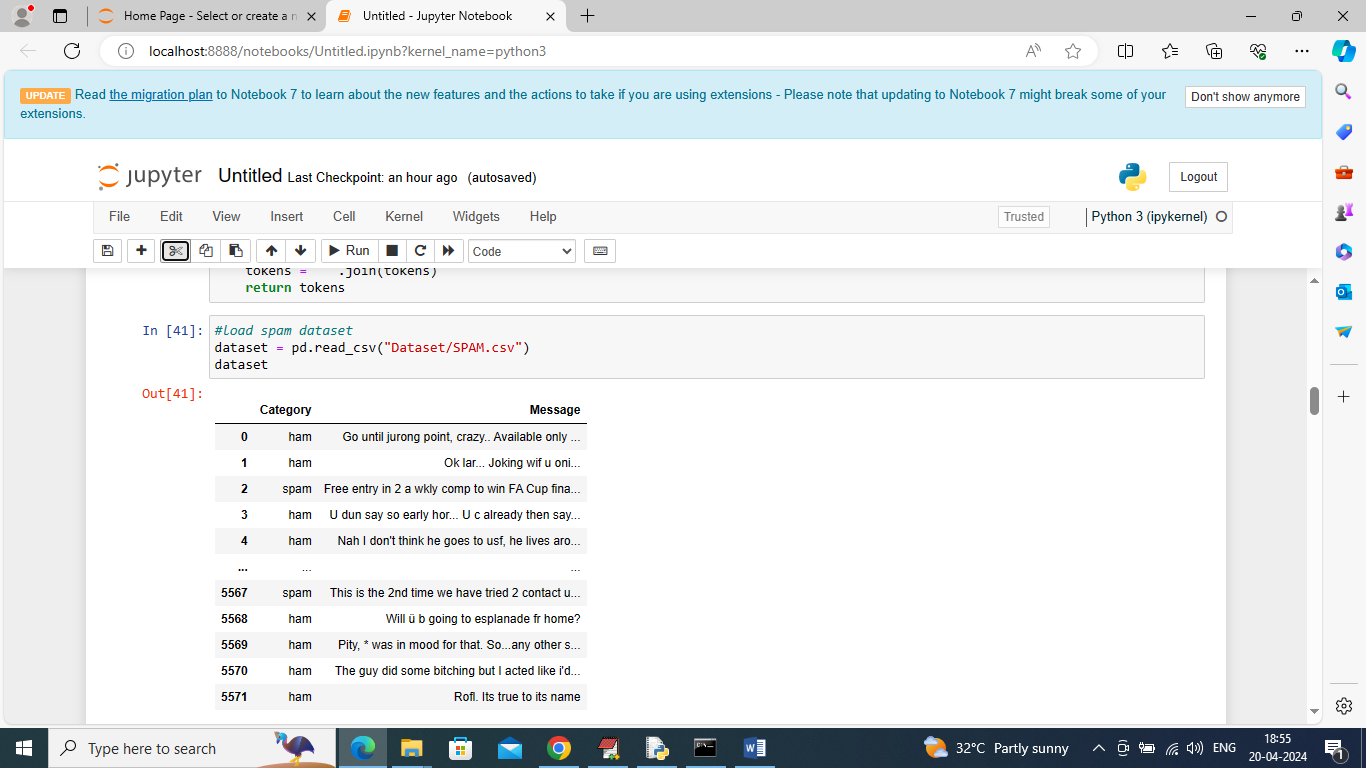
To train LSTM we have utilized SMS SPAM dataset given to implement this task we have implemented this project using JUPYTER tool and below are the code and output screens



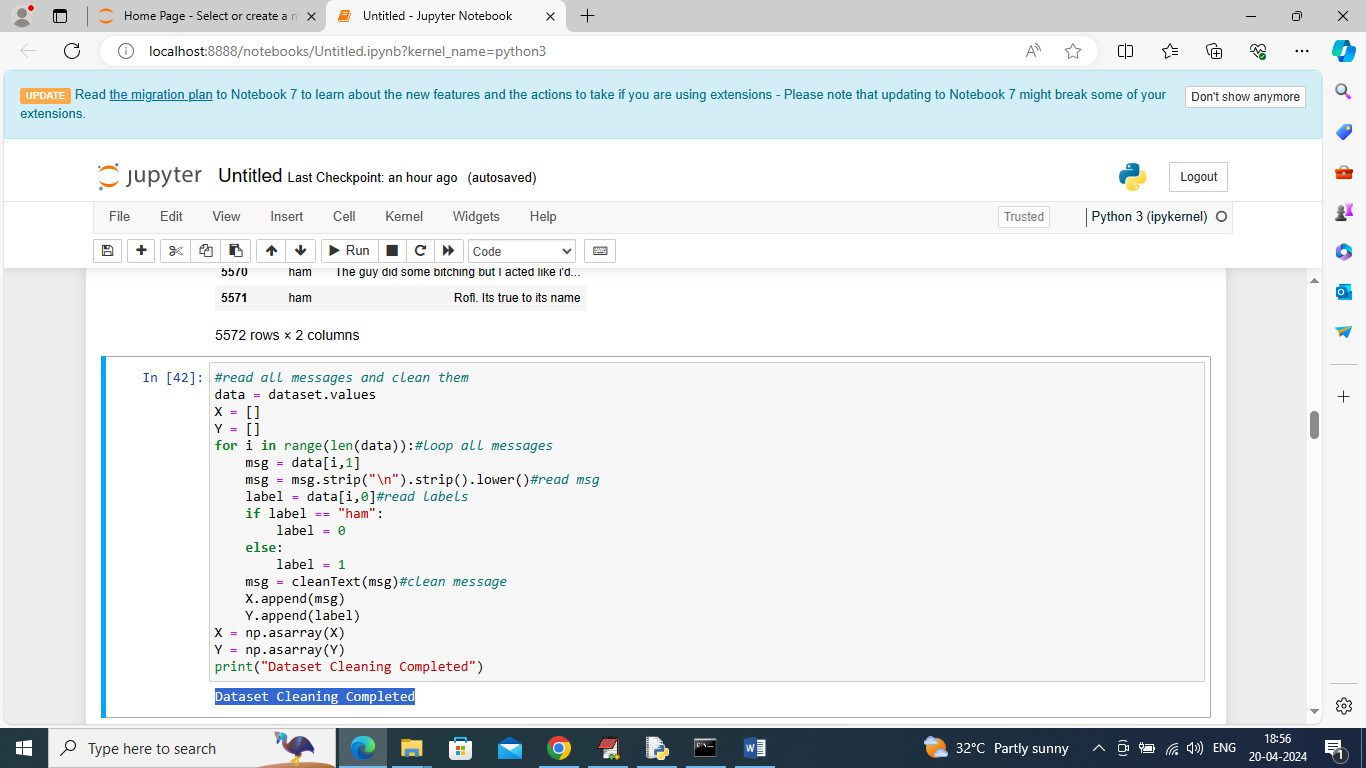
Above screen shots showing loading of required classes and packages



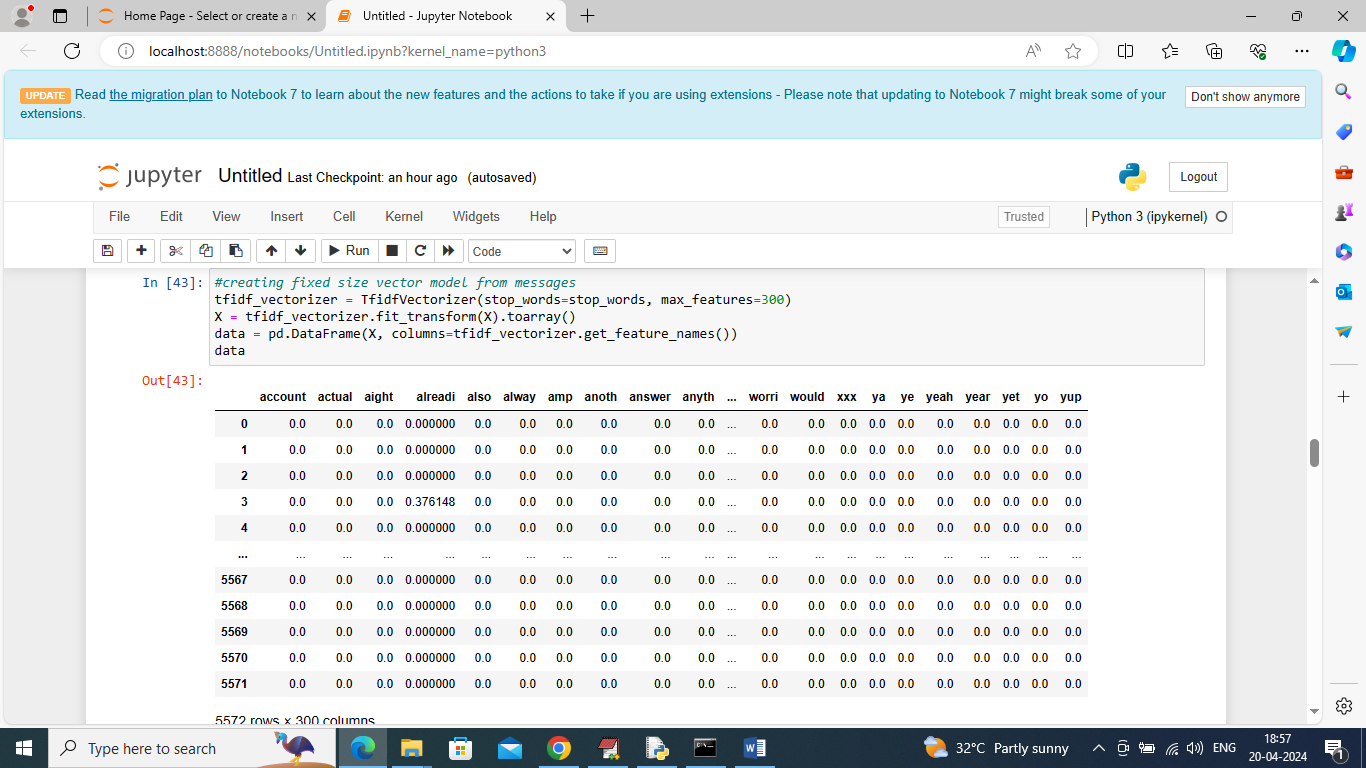
in above page we have utilized classes from NLTK package to remove stop words, stemming and lemmatization and in above screen cleaning text message with this classes



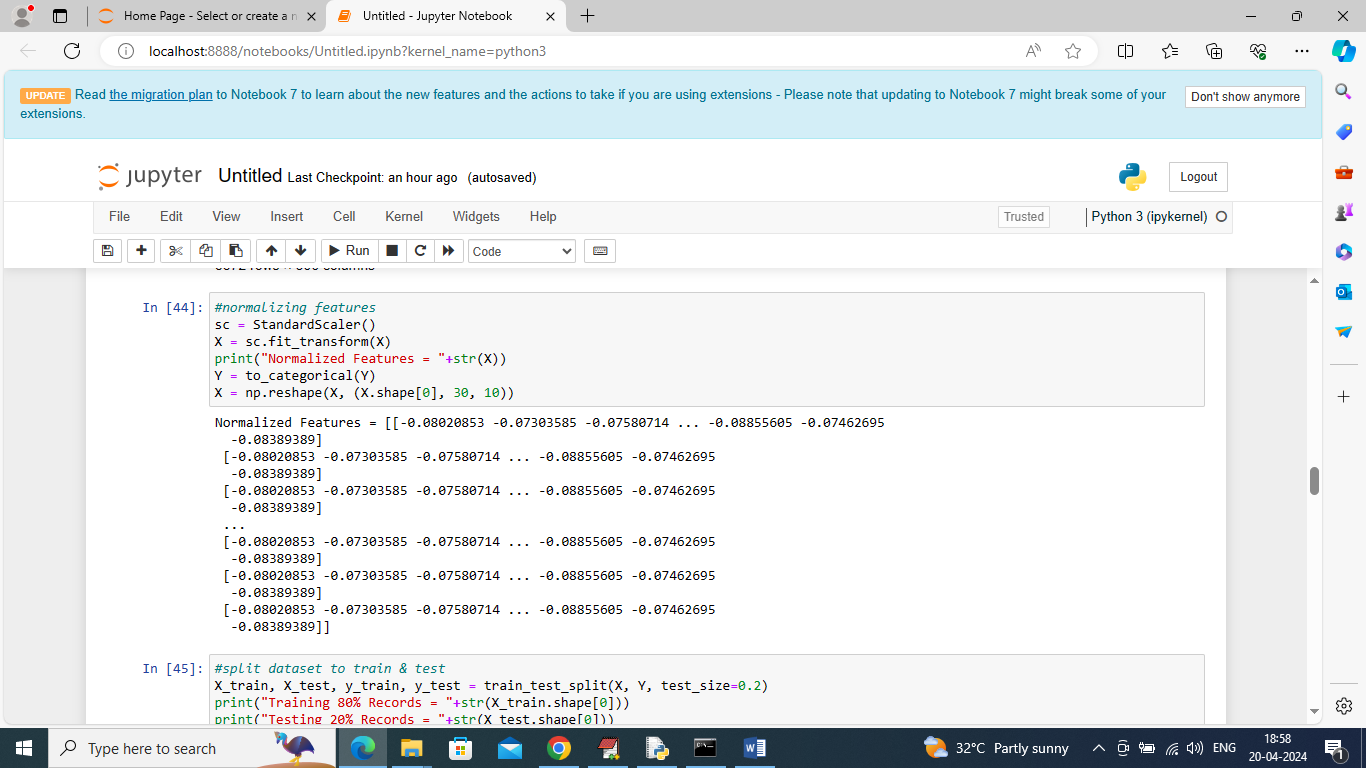
In above screen loading and displaying dataset values



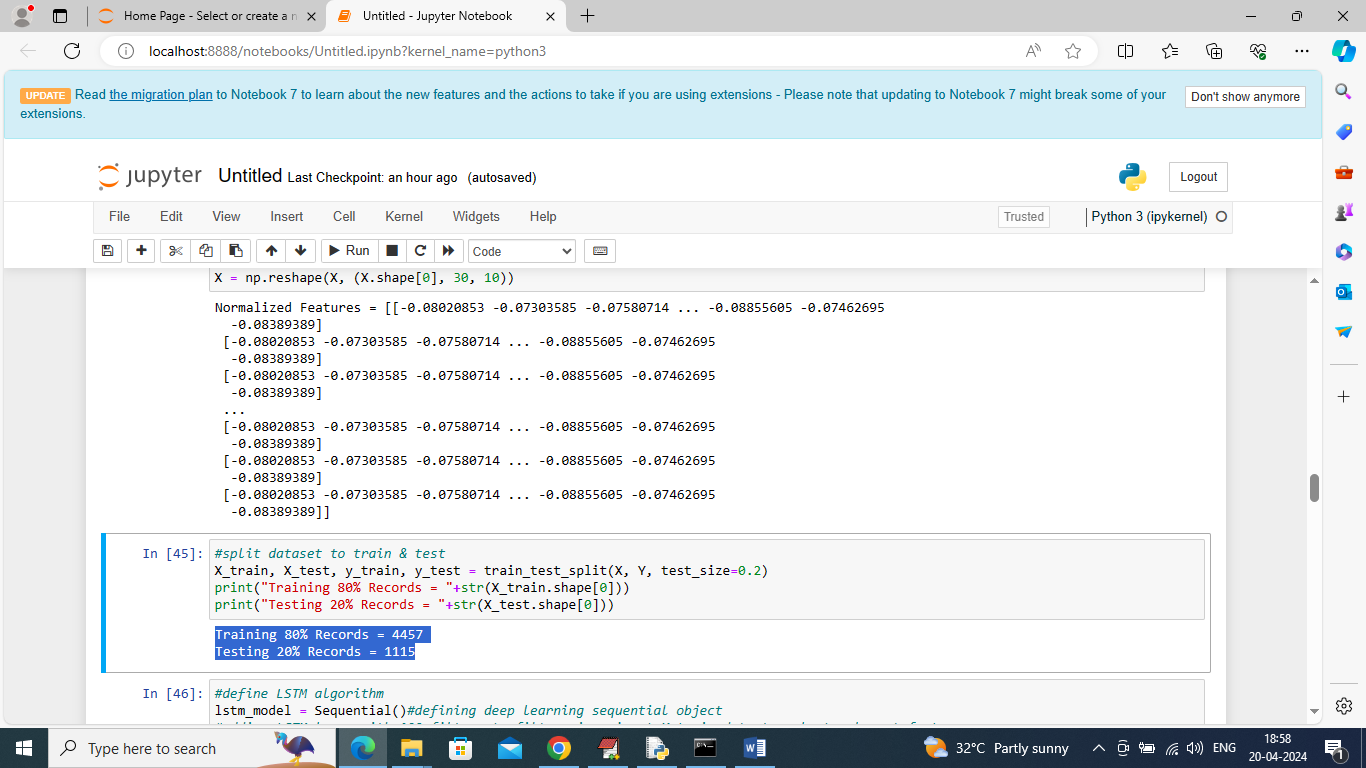
In above screen looping all messages and then calling CLEAN function to remove stop words, stemming and lemmatization will be applied and then create X and Y training array



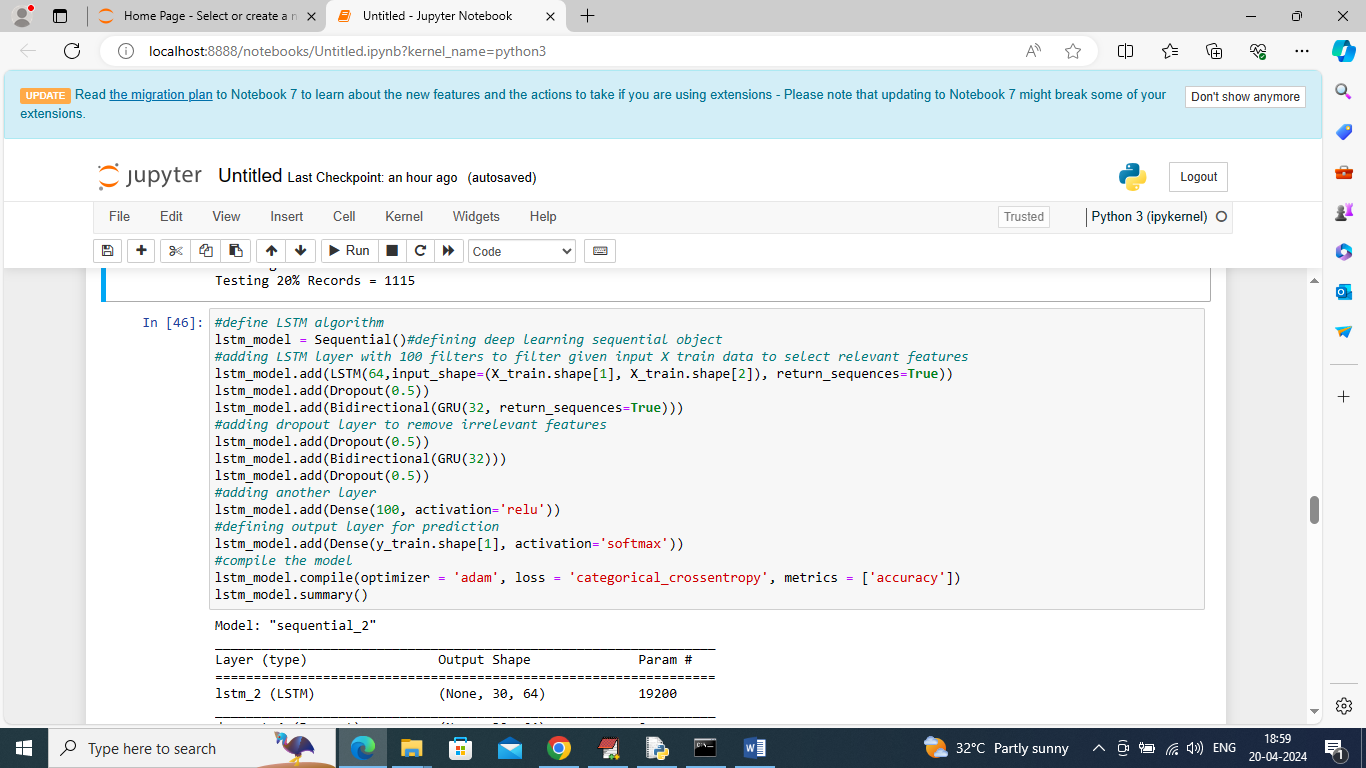
In above screen using vector class we have converted all text into fixed size numeric vector and in above vector all column contains word NAMES and remaining rows contains average occurrence of that column words



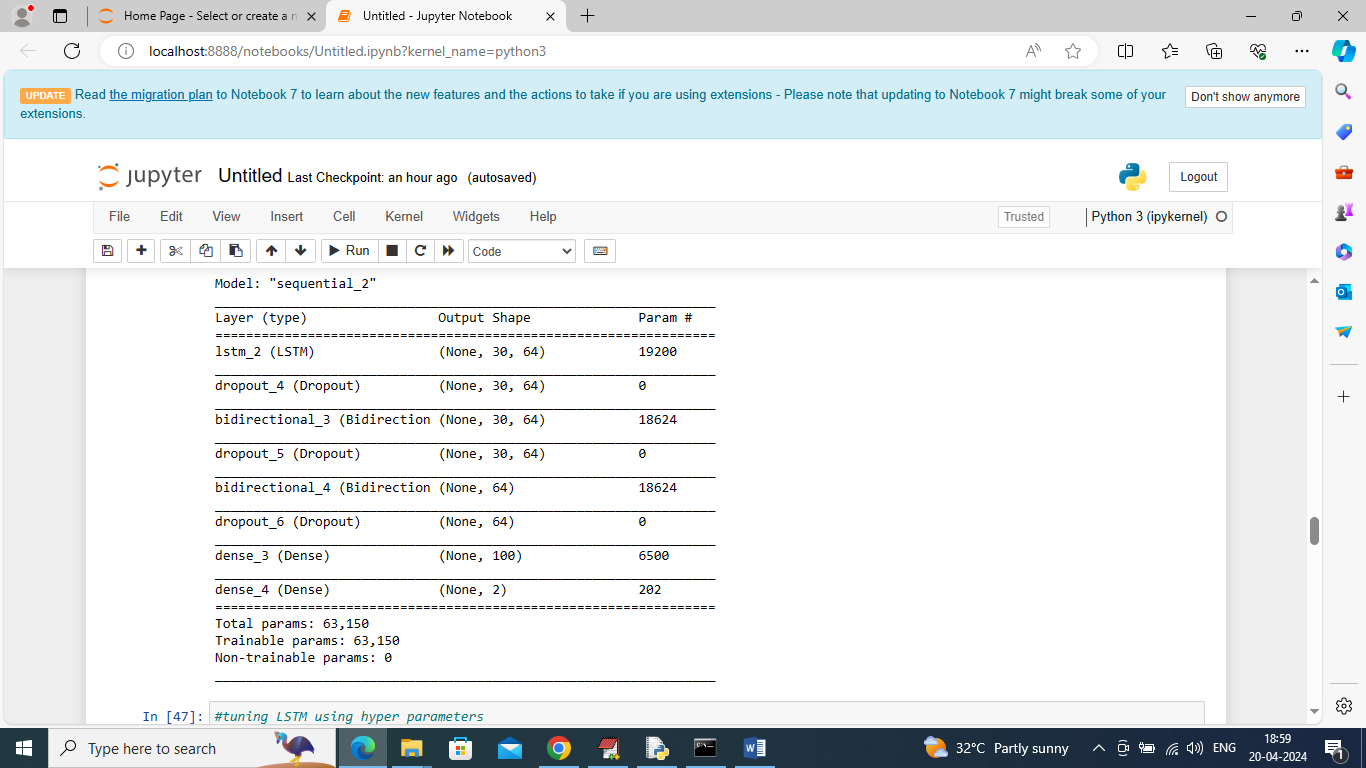
In above screen normalizing entire numeric vector and then displaying normalize values



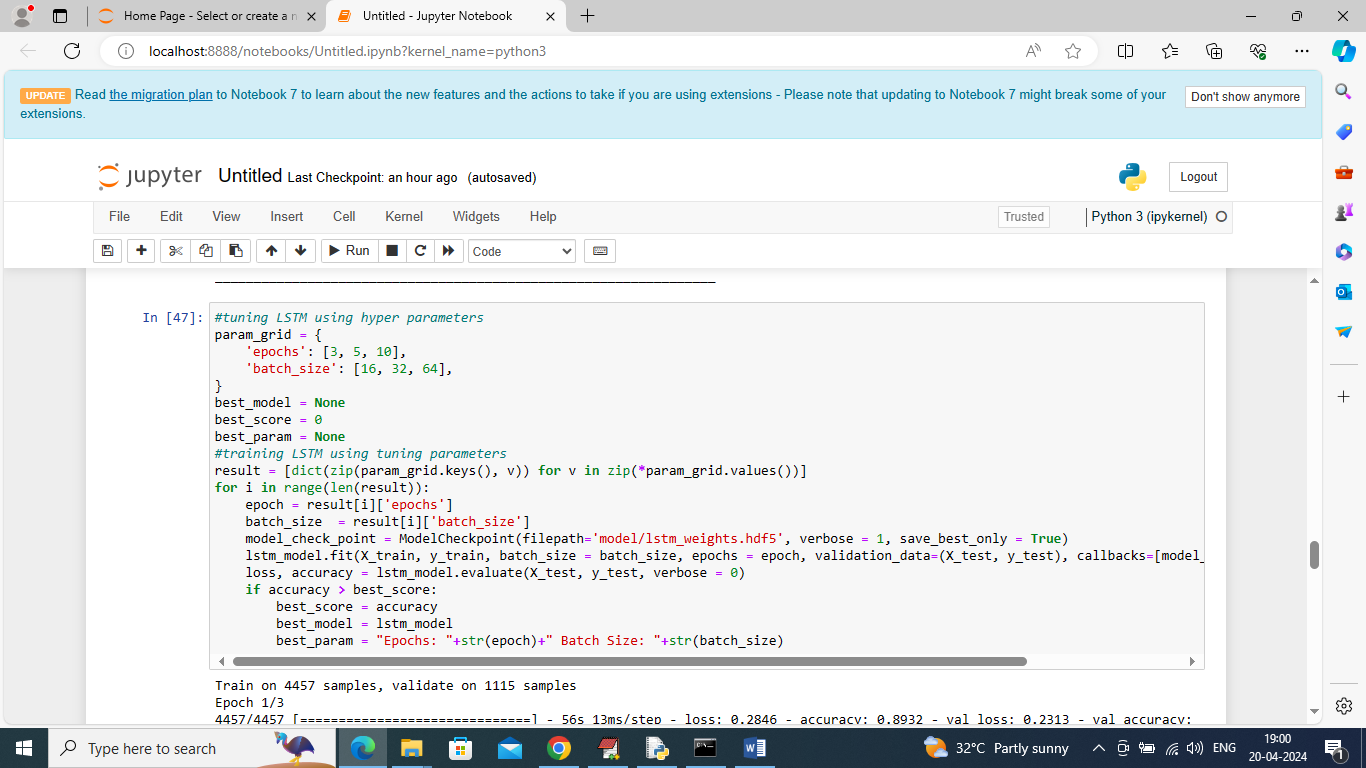
In above page splitting dataset into train and test where application using 80% dataset messages for training and 20% for testing



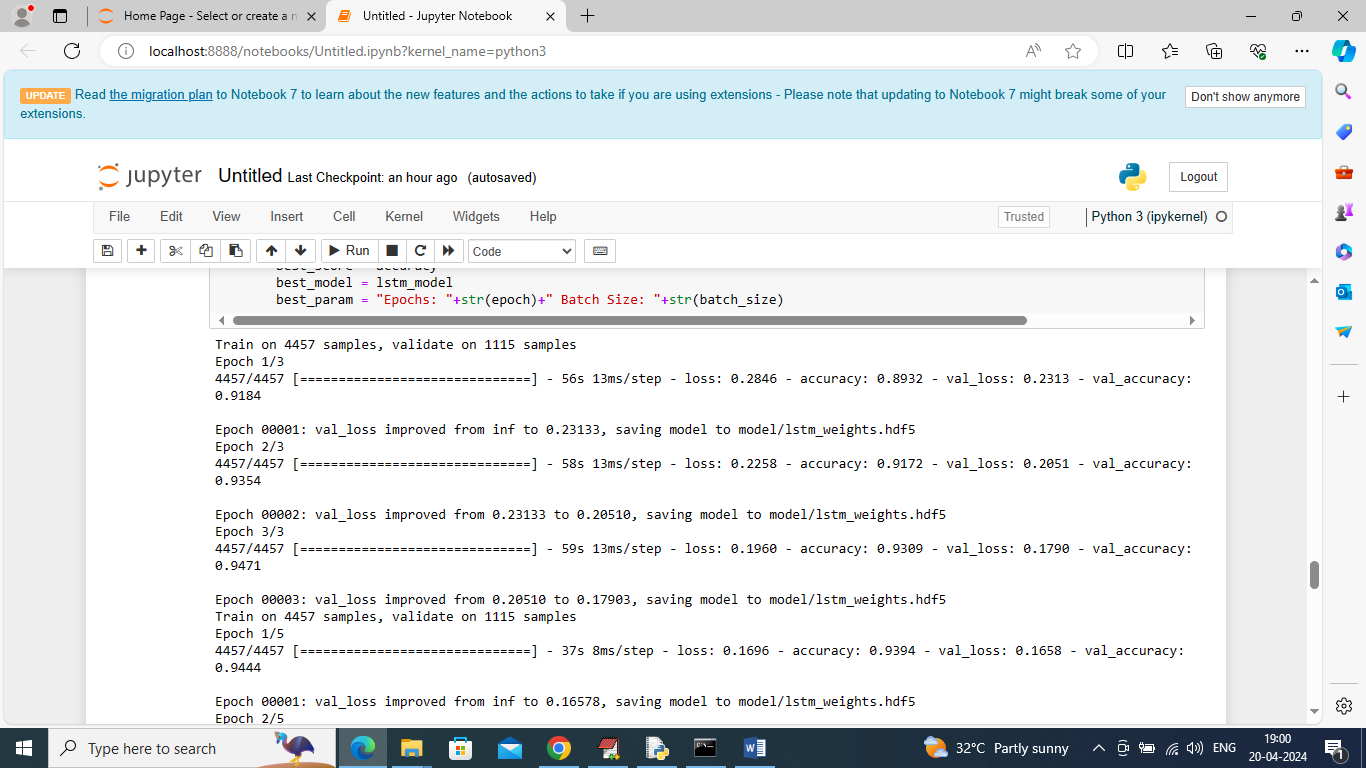
In above screen defining LSTM architecture and below is the summary of above model



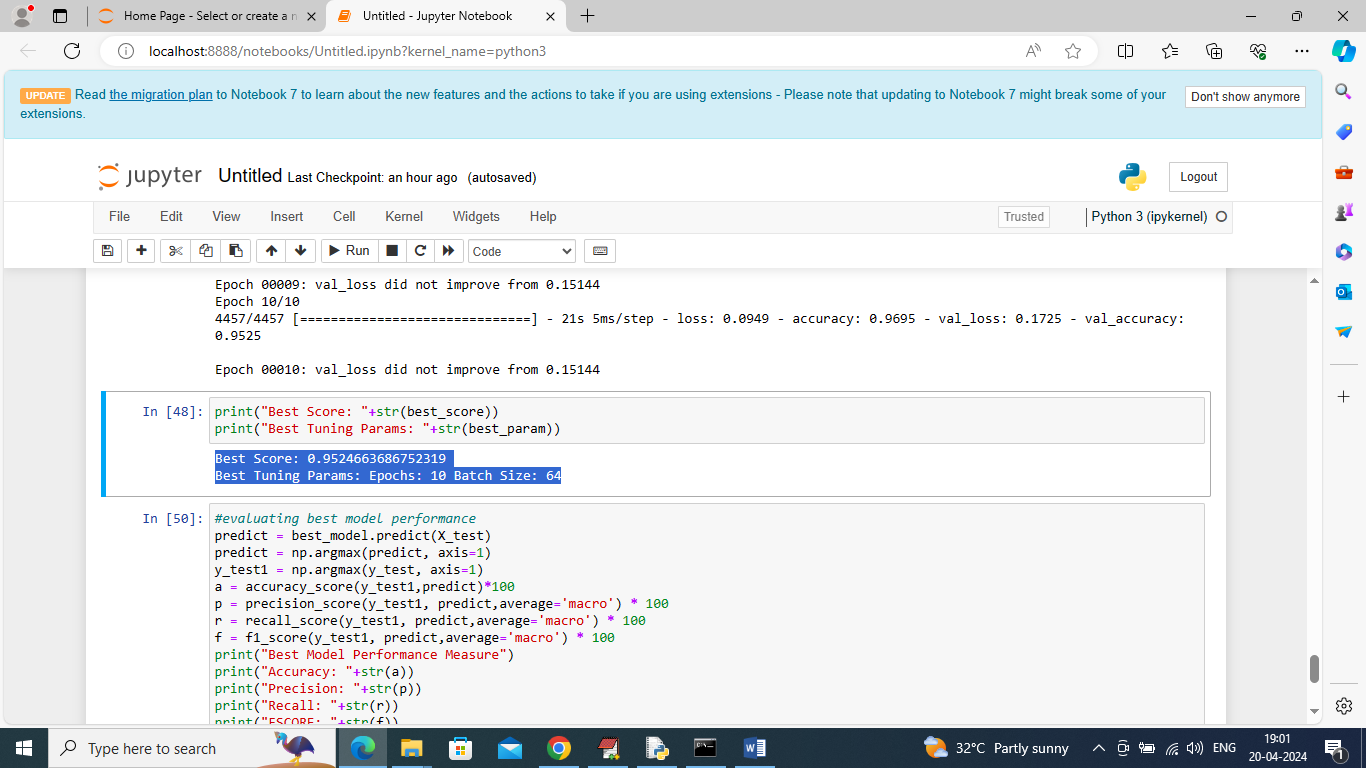
In above screen can see LSTM model summary



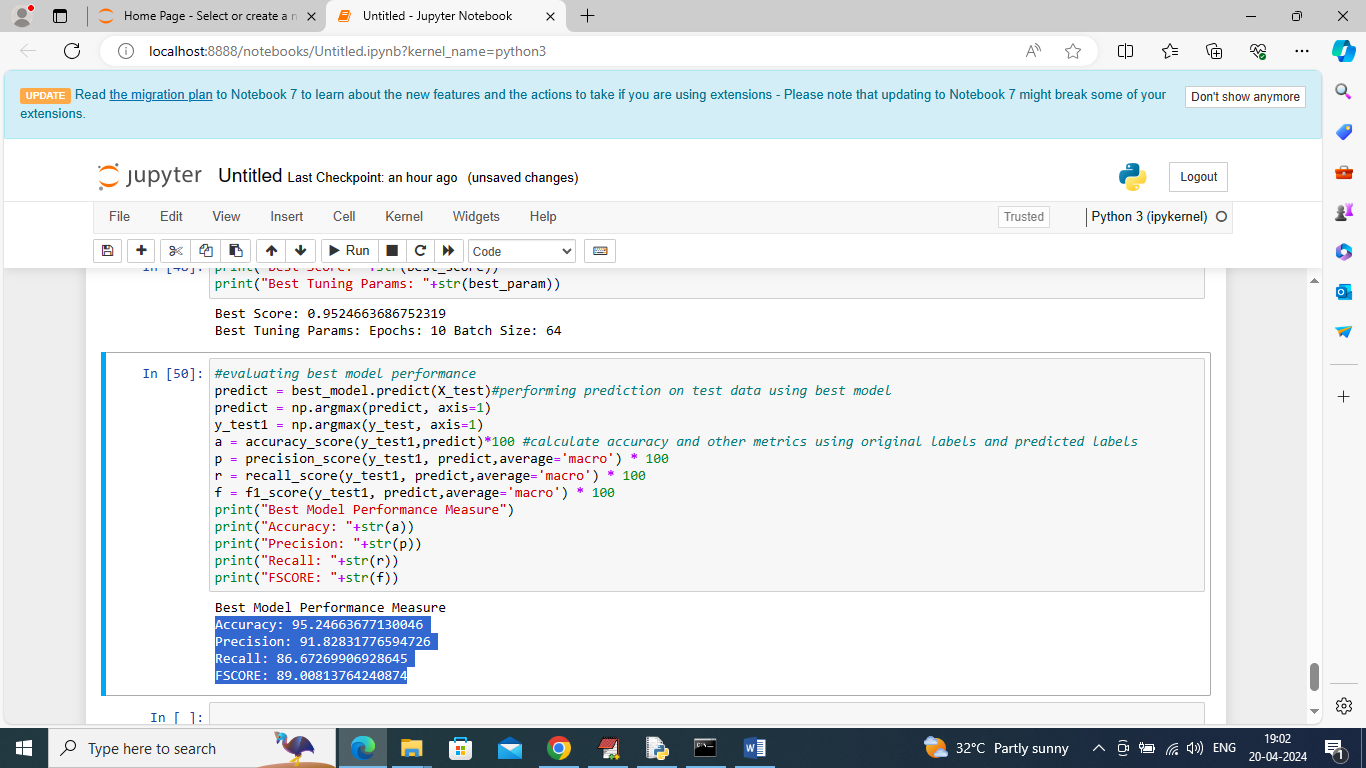
In above screen training LSTM model by employing tuning parameters and while training will get below output



In above screen LSTM starts training as per tuning parameters and after all parameters will get below best score and parameters values



In above screen can see best model score and tuning parameters and now we are performing prediction on test data using BEST MODEL



In above page in blue colour text can see accuracy, precision, recall and FCSORE of best model predicted on unknown 20% test data and this model able to classify SPAM messages with an accuracy of over 95%