**Result:**

**1] Model Accuracy result:**

To evaluate our model, we have trained and tested our model on the amazon dataset. To compare the accuracy of the model, we have made a deep learning model and compared it with machine learning models. Firstly, we have generated the RNN model with LSTM (Long Short Term Memory) and got the training and testing results shown in figure 1 and figure 2.

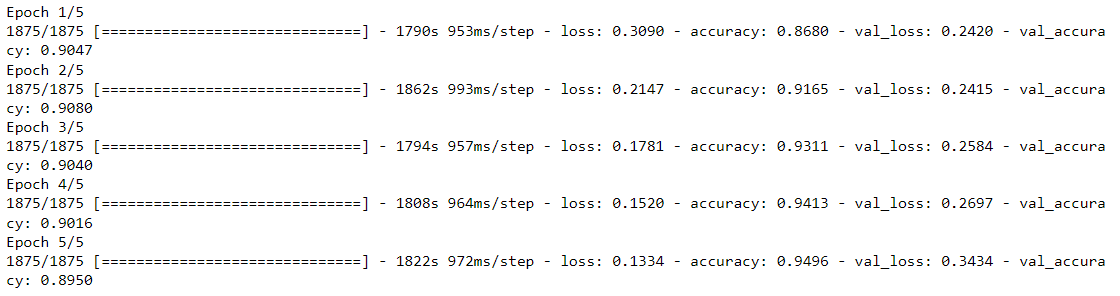


Figure 1



Figure 2

And to compare RNN model accuracy, we have generated 4 different machine learning models namely Multinomial Naïve Bayse, Decision Tree, Logistic Regression, and Random Forest and we got the accuracy score of 50.22%, 54.88%, 50.26%, and 54.06%, respectively. For the evaluation, we have generated a graph that contains all 5 model results as shown in figure 3. It can be clearly seen that RNN a deep learning model performs far better than any other machine learning model.

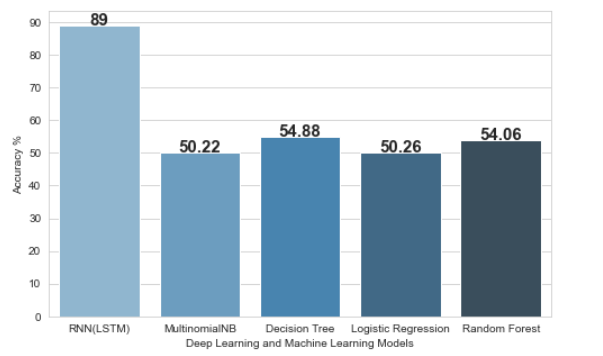


Figure 3

**2] Real time review analysis:**

To move further in our project, we have established a real-time review analysis model that can predict sentiment by fetching input provided by the users. This model can work on the text, emoticons as well as on emojis. As shown in the below figures, the model can also predict sentiment from emoticons and emojis with the best accuracy.

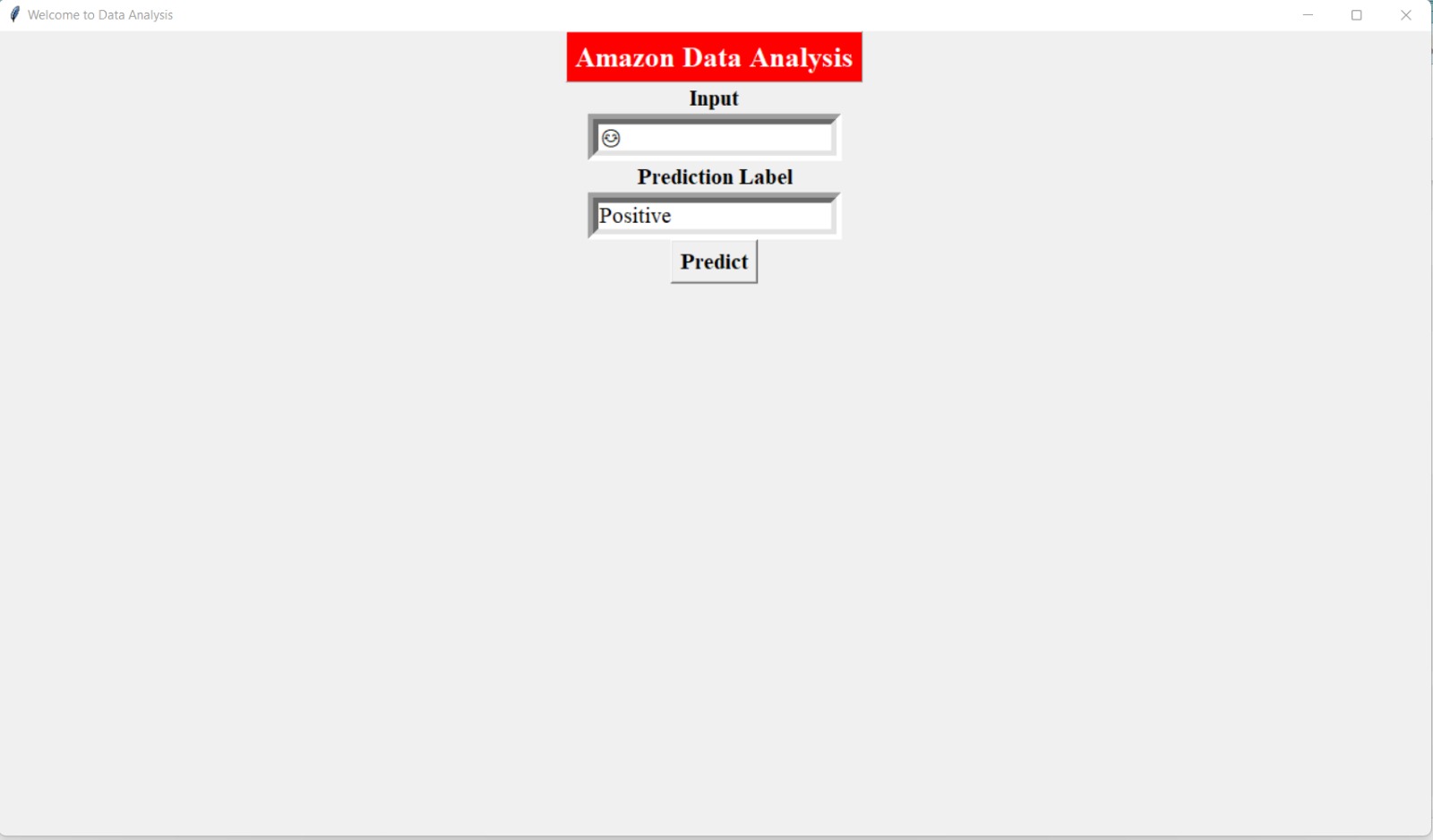






**3] GUI:**

Lastly, we have incorporated our model with Graphical User Interface that makes user’s task far much better and understandable. This GUI contains two fields in which first to fetch user’s input and gives output in second field as shown in the figure.



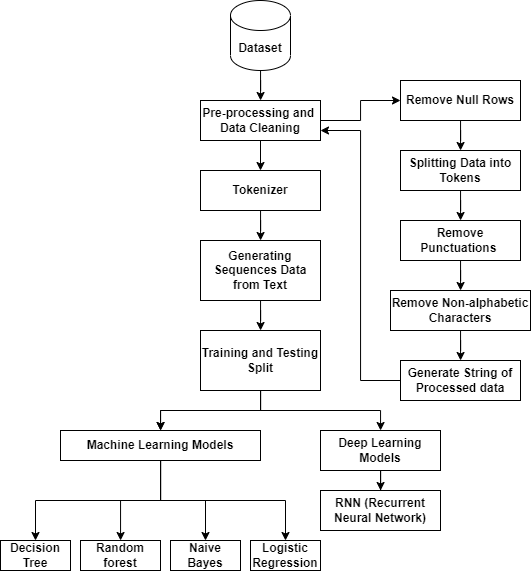
**Methodology:**

In this section, we will present the methodology and techniques used in classifying the Amazon product that is used by most of the researchers in the field of sentiment analysis. And here we also propose a potential solution for the classification of sentiment as there are some flaws in the current research. We have targeted to eliminate those flaws. We choose to try different machine learning approaches to get better results for accurate sentiments of the document and review. The research follows the steps described below.

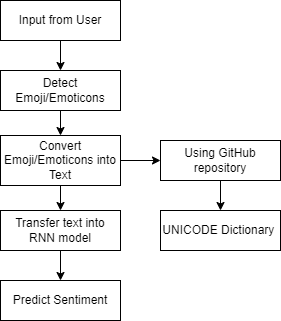
Initially, we performed a sentiment analysis task on the Consumer Reviews of Amazon Products dataset [1] which is openly available on Kaggle and we got good accuracy with that dataset. But when we perform real-time review analysis on those trained models we got only positive sentiment for any input review. This happens because of imbalanced data that contains almost 24k positive labeled data and only 400 negative labeled data. So we have changed our dataset and performed operations on the Amazon Reviews for Sentiment Analysis [2] dataset that contains 4,00,000 of balanced data with 50% positive reviews and 50% negative reviews.

Initially, we mapped the label field with its appropriate sentiment and then we preprocessed the raw data to get rid of null values, punctuations, stop words, apostrophes, and hyphens. After that, we tokenized the review string and generated the preprocessed string from the tokens. After that, we factorized the sentiment field to map it with reviews. For better accuracy and performance, we converted the text data into sequence data and we generated a padded sequence for each review that can boost the accuracy of the model.

For the training and testing part, we have implemented a deep learning model that is LSTM (Long Short Term Memory) which is a type of Recurrent Neural Network. In which, we made an embedding layer with 32 embedding vector lengths and 400 input sizes. We have also considered a spatial dropout of 0.25 and a dense layer with a “Sigmoid” activation function. In the end, we compiled it with “Adam” optimizer and calculated the accuracy for the training and testing set of the data. For the evaluation of the deep learning model, we have passed the same training and testing data to the four different machine learning models which are multinomial naïve bayse, decision tree, logistic regression, and random forest.



From the previous experiment, we found that the deep learning model gives better results than any other machine learning model so for the future task we have considered the LSTM model. Our main focus of the project is sentiment analysis from emojis and emoticons with real-time reviews. So we have considered a GitHub repository [3] that contains a dictionary of emoticons and emojis that helped us to convert the emoticons and emojis into their appropriate text. In this, we fetch the input from the user and predict the sentiment by passing the review into the pre-trained model. And if the review contains any emoji or emoticon then firstly, we convert it into text and then pass it into the model so that the user can get accurate sentiment for text as well as for emoji.



Lastly, to increase the user’s efficiency and understandability we have made a Graphical user interface that fetches the user input and predicts the sentiment by passing that input into the model and printing the result into another field which helps the user to understand the process easily.

References:

[1] DATAFINITI, “Consumer Reviews of Amazon Products”

[2]

[3]