**INTERNSHIP: PROJECT REPORT**

------------------------------------------------------------------------------------------------------------------------------------------

|  |  |
| --- | --- |
| Internship Project Title | RIO-45: AUTOMATE DETECTION OF DIFFERENT EMOTIONS FROM TEXTUAL COMMENTS AND FEEDBACK |
| Project Title | DEEP LEARNING CLASSIFICATION ALGORITHM TO PREDIT DIFFERENT TYPES OF EMOTIONS FROM TEXTUAL COMMENTS AND FEEBACKS |
| Name of the Company | TATA CONSULTANCY SERVICES |
| Name of the Industry Mentor | MR.DEBASHIS ROY |
| Name of the Institute | MALLA REDDY COLLEGE OF ENGINEERING AND TECHNOLOGY |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Start Date | End Date | Total Effort (hrs.) | Project Environment | Tools used |
| 20-03-2022 | 24-03-2022 | 20 | Google Colab, Visual Studio code(jupyter notebook),Windows 10 | Python 3(Tensorflow, Keras) |

|  |
| --- |
| PROJECT SYNOPSIS:  Emotion analysis is the process of identifying and analyzing the underlying emotions expressed in textual data.. Emotion analytics can extract the text data from multiple sources to analyze the subjective information and understand the emotions behind. Emotion analysis is focus on targeting to automatically extract the emotional states of users from their reviews (e.g., blogs, tweets, post, etc.) in social networks. In other words, if you want to maximize purchases, maximize peoples’ emotions about the product. The emotion analysis based on bipartition and tripartition, it can not fully express the complex inner world of people, which ignores subtle emotional changes, nor it can describes user’s psychological state more comprehensively. Emotional analysis is also called fine-grained sentiment analysis. Thus, sentiment analysis can regarded as a bipartition oriented emotion analysis. From the perspective of human psychology, emotional analysis can characterize emotional attitude of people in multi-dimensional view, based on the existing coarse-grained sentiment analysis. The overall customer experience make a review of the product. It also categorize feedback of movies, products, ect.. Hence, developing a tool for automating Emotion Analysis can be very beneficial for various companies such as online shopping sites, online food order websites ,etc |
| ASSUMPTIONS:  I had imported two datasets training data and testing data of a reviews on a product. It contain a shape of (11237, 2). It had emotion and Text columns. I had been imported from github repository. |
| PROJECT DIAGRAMS:  MODEL: |
| ALGORITHM:  Step 1:Import the required libraries from tensorflow and keras. And also imported datasets from github.  Step 2: After, I had done Exploratory data analysis and Feature Engineering on text data.  Step 3: Form some visual analysis using seaborn and matplotlib libraries.  Step 4:Imported wiki data of news papers for Word embedding.  Step 5: Build the model with 3 layers.   1. The first layer is an Embedding layer. An embedding layer stores one vector per word. When called, it converts the sequences of word indices to sequences of vectors. These vectors are trainable. After training (on enough data), words with similar meanings often have similar vectors. The resulting dimensions are: (batch, sequence, embedding). 2. The Second Layer is LSTM with dropout 0.2, recurrent drop of 0.2.This increase the Forward and backward propagation of the RNN layer then combines the output. LSTM will store more data and it make more efficient output. As well as a GRU it is similar to LSTM but difference is it follow binary classification. It build on If-Else condition. 3. This index-lookup is much more efficient than the equivalent operation of passing a one-hot encoded vector through a *tf.keras.layers.Dense* layer. A recurrent neural network (RNN) processes sequence input by iterating through the elements. RNNs pass the outputs from one timestep to their input. Softmax function is used for layer   Step7: Compile the model with the loss function and optimizer. Here, Loss function is “categorical\_crossentropy” since it has the Multi class classification.So it measures the distance between the probability distributions.  Step 8:Fit the model train data, with epochs =12 , making a validation with test data.  Here  Epoch 1/12  62/62 [==============================] - 397s 6s/step - loss: 1.4770 - accuracy: 0.3571 - val\_loss: 1.3307 - val\_accuracy: 0.4400  Epoch 2/12  62/62 [==============================] - 390s 6s/step - loss: 1.2275 - accuracy: 0.5000 - val\_loss: 1.1320 - val\_accuracy: 0.5511  Epoch 3/12  62/62 [==============================] - 391s 6s/step - loss: 1.0278 - accuracy: 0.6075 - val\_loss: 0.9422 - val\_accuracy: 0.6466  Epoch 4/12  62/62 [==============================] - 387s 6s/step - loss: 0.8569 - accuracy: 0.6865 - val\_loss: 0.8332 - val\_accuracy: 0.6905  Epoch 5/12  62/62 [==============================] - 387s 6s/step - loss: 0.7788 - accuracy: 0.7153 - val\_loss: 0.8103 - val\_accuracy: 0.7073  Epoch 6/12  62/62 [==============================] - 377s 6s/step - loss: 0.7495 - accuracy: 0.7255 - val\_loss: 0.7929 - val\_accuracy: 0.7056  Epoch 7/12  62/62 [==============================] - 393s 6s/step - loss: 0.7163 - accuracy: 0.7409 - val\_loss: 0.7621 - val\_accuracy: 0.7250  Epoch 8/12  62/62 [==============================] - 391s 6s/step - loss: 0.6968 - accuracy: 0.7470 - val\_loss: 0.7641 - val\_accuracy: 0.7283  Epoch 9/12  62/62 [==============================] - 380s 6s/step - loss: 0.6820 - accuracy: 0.7561 - val\_loss: 0.7506 - val\_accuracy: 0.7265  Epoch 10/12  62/62 [==============================] - 388s 6s/step - loss: 0.6672 - accuracy: 0.7590 - val\_loss: 0.7492 - val\_accuracy: 0.7280  Epoch 11/12  62/62 [==============================] - 383s 6s/step - loss: 0.6531 - accuracy: 0.7654 - val\_loss: 0.7421 - val\_accuracy: 0.7297  Epoch 12/12  62/62 [==============================] - 387s 6s/step - loss: 0.6392 - accuracy: 0.7750 - val\_loss: 0.7392 - val\_accuracy: 0.7327  Step 8: Predict Emotion with sample data.  Example:  Message: ['There was a hairline scratch on the screen which is not clearly visible directly. When we tried contacting Amazon for the replacement, they asked us to connect to apple and when connected to apple, they asked to connect with Amazon. Blame game begins and we end up paying the price. Awful service. I am going to tag on the social media both Amazon and apple. Such a big scam']  predicted: anger (0.44 seconds) |
| EXPECTATIONS CONSIDERED:  Accuracy & Loss Function values vary slightly on every compilation of the model. This may give a slight difference in the prediction value which may impact the neutral network. Padding data is possible for large input data also. |
| ENHANCEMENT SCOPE:  Increasing the accuracy is the biggest enhancement scope. This can be achieved by adjusting the LSTM layer with dropouts. It achieved maximum output for one LSTM layer. |
| LINKS TO CODE AND EXCUTABLE FILE:  Link of project work (notebook)  [Colab\_Link](https://colab.research.google.com/drive/1lqtWceHe_QgpAAiyeHE6jEycP2whkcJa?usp=sharing) |