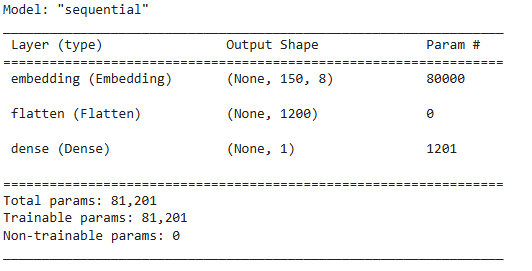
**Summary**

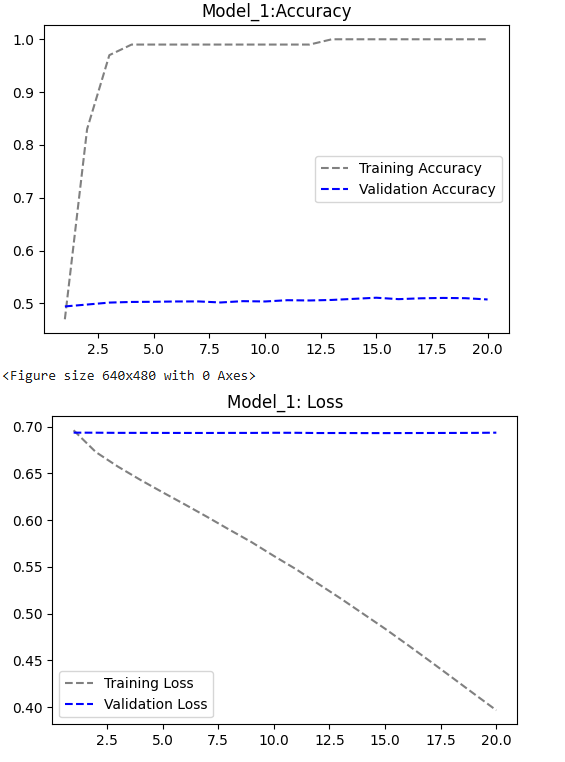
First, I have Loaded the IMDB Dataset and taken maximum feature number of words to 10000 and cutoff text length to 150 words. Load all the required libraries and after this step used an embedded layer and classifier on the IMDB dataset.



The above picture shows the sequential model that I have obtained with number of parameters. As embedding is equal to vocabulary size i.e., input\*output = 10,000\*8 = 80000

Dense is the maxlen i.e., input\_length\*output +1 = 300\*8+1 = 2.401

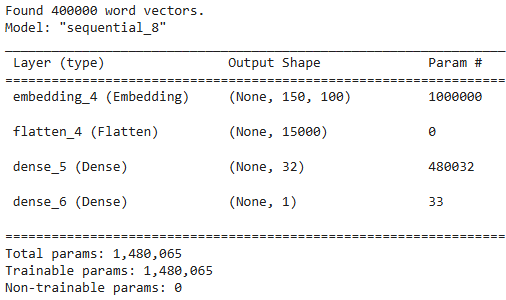
Therefore, Total Parameters = 82.401



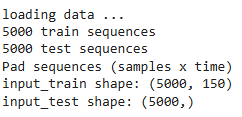
The above graph has been obtained with an validation accuracy of 51% and validation loss of 69%.

Now used a pretrained word embeddings and tokenizing the data and used GloVe, or Global Vectors for Word Representation, is an unsupervised learning algorithm for generating vector representations of words based on their co-occurrence statistics in large text corporation. We are getting the data from ai.stanford.edu

Loaded pretrained word embedding into the embedding layer and evaluated the model.



Now consider an LSTM example an load the IMDB dataset



The below table shows the accuracy and loss of each model.

|  |  |  |
| --- | --- | --- |
| Model | Accuracy | Loss |
| Embedded Layer | 51% | 69% |
| Glove Embedding | 50% | 99% |
| LSTM | 81% | 55% |