Saarthi.AI (Extracting Labels NLP) :

We were given the task of extracting the labels from a text transcription .

To be exact we were given a text and we were to divide it into three labels

Of action , object and location. We had a total of 11000 instances of training dataset

And 3000 instances of validation dataset.

I followed the following steps for reaching completing the task .

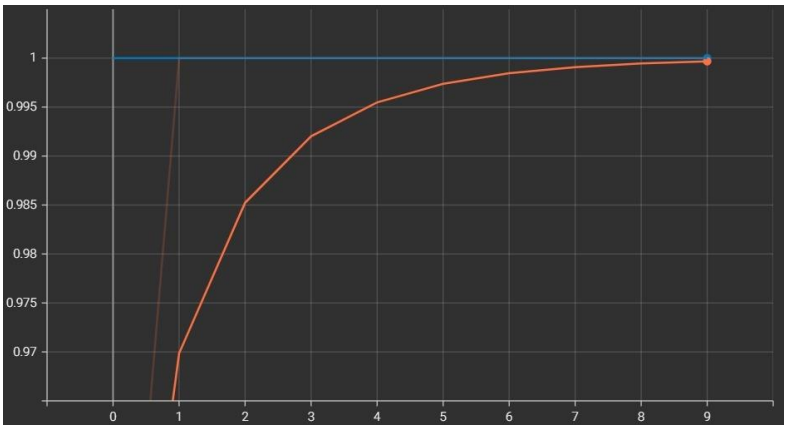
1. First of all I visualized the data using pandas , I got the info about the number of Labels such as there were 6 labels for action , 14 labels for object and 4 labels for Location.
2. Further after analyzing the data , I found that I don't need to apply much text Cleansing and can work with the given data.
   Thus I applied only lowercase function for changing the given transcription into lowercase.
3. After going through various models for our task of extracting labels I decided to Go with the universal sentence encoder model is a widely known sentence encoder model . It had a better performance than other models like bert model etc.
4. Further I trained the model for about 10 epochs with a batch size of 64 for training purpose And which can be varied as per our requirement.
5. I used tensorboard for logging my results and used json config file.

I received an astonishing result of about 100% accuracy on both the train and validation data.

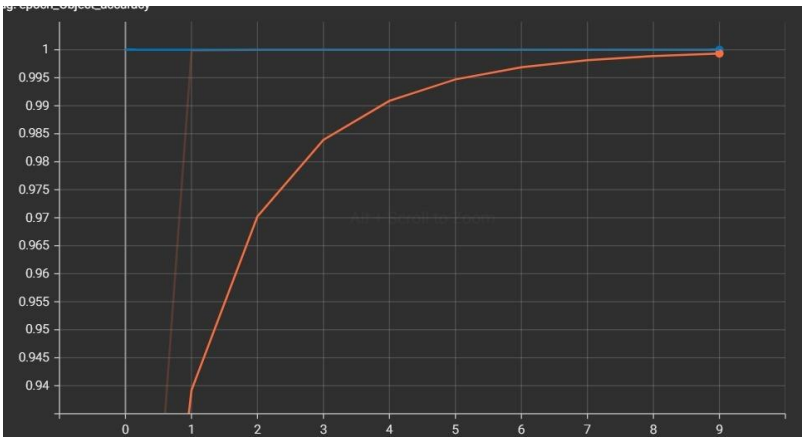
Plots :

Training accuracy

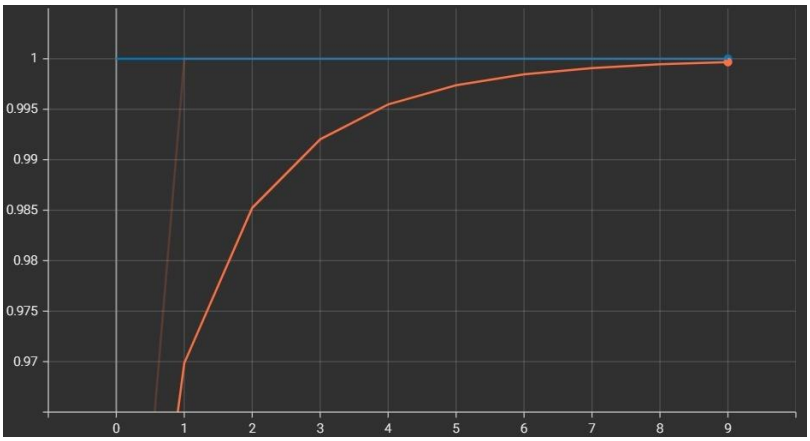
Validations accuracy -



Obtained plot for the action



Obtained plot for object label



Obtained plot for location label