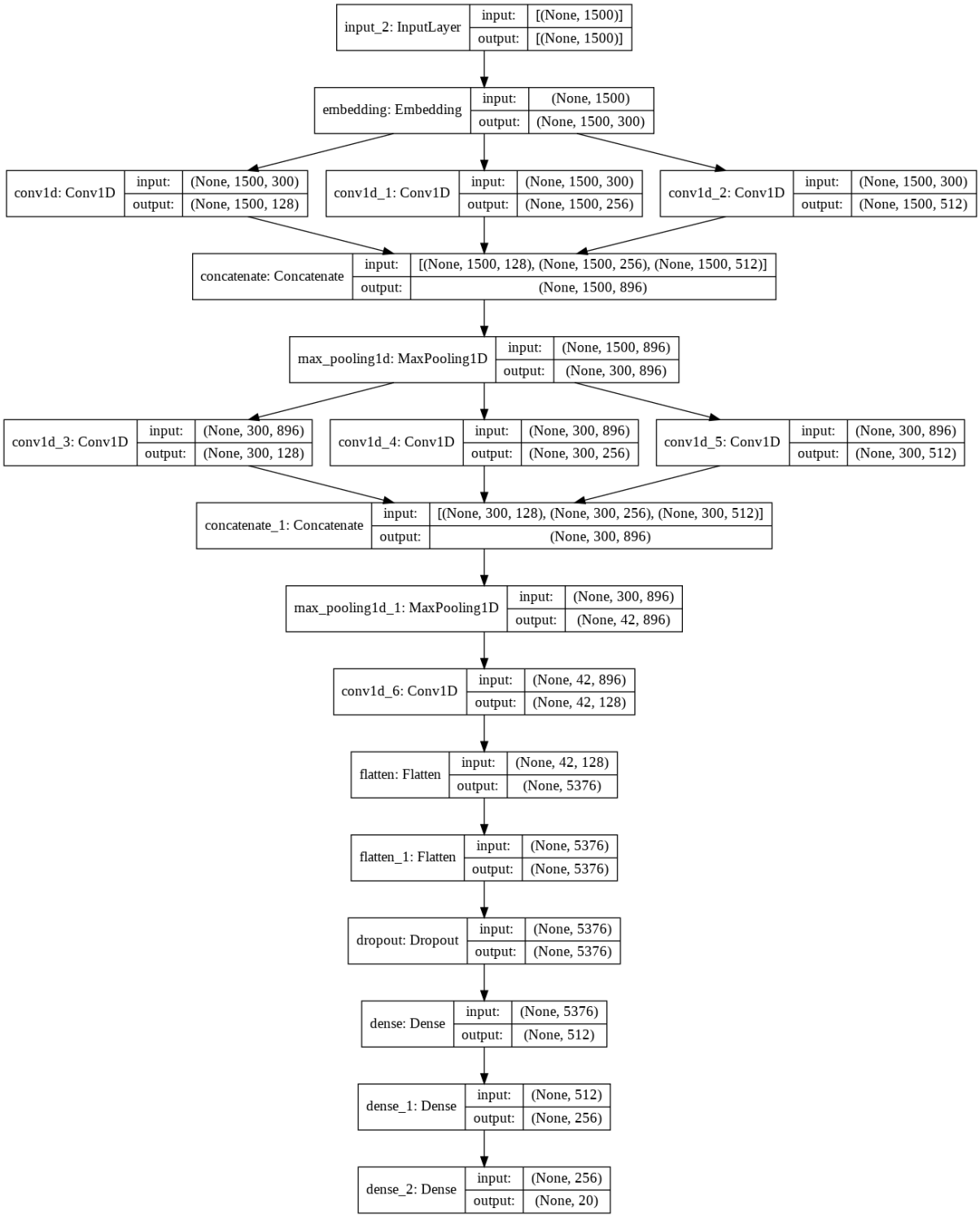
Observations from CNN assignment

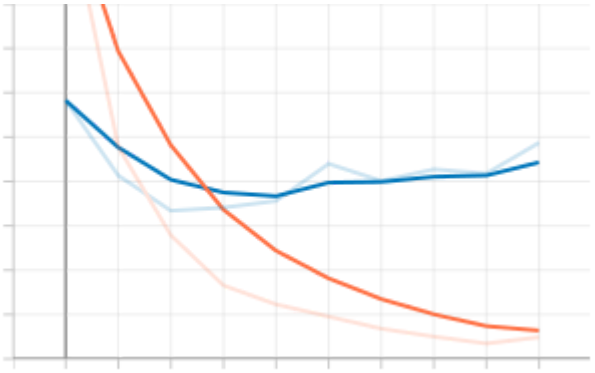
- Model 1:

1. Model architecture:

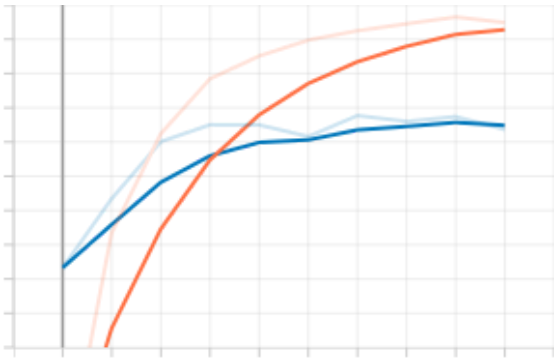


2. Loss and accuracy plots from tensorboard:

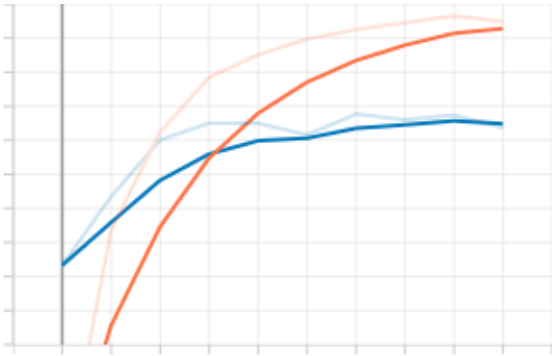
Epoch loss



Epoch accuracy

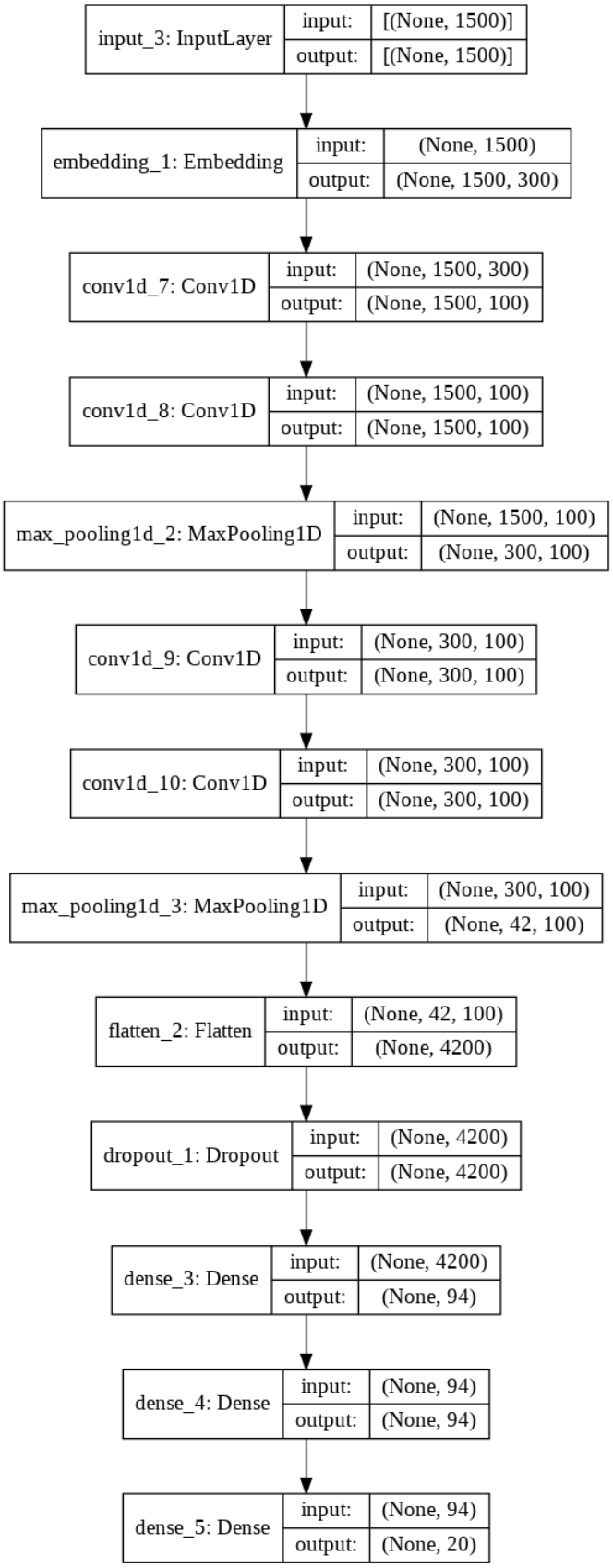


Epoch F1

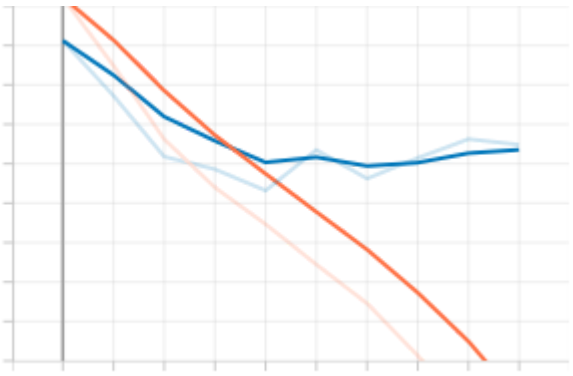


- Model 2:

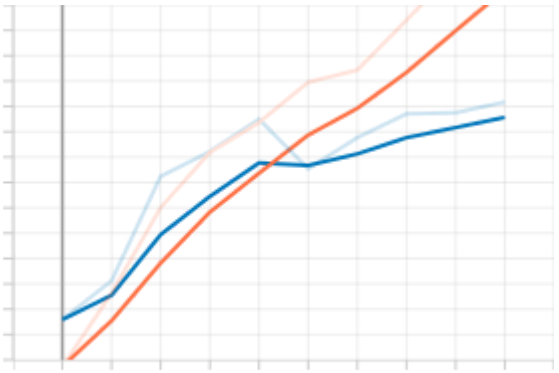
1. Model architecture:



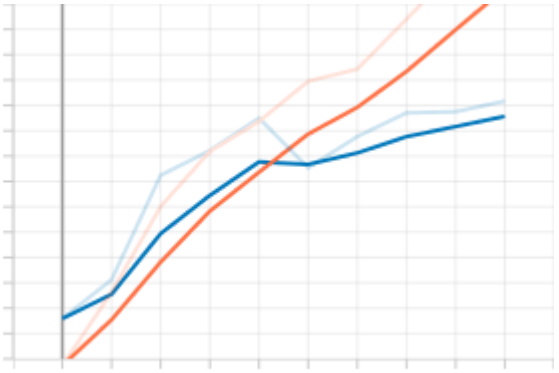
2. Loss and accuracy plots from tensorboard:
   Epoch loss



Epoch accuracy



Epoch F1



- About models:

    - This is a Text Classification model used with CNN on the famous 20 Newsgroups dataset. There are a total of 2 models tried with varied architectures
    - In the first model, a word embedding layer is used with the help of pre- trained 300D glove vectors followed by 1D Convolution and Maxpool layers.
    - Whereas in Model 2, character embeddings are used along with the aforementioned layers but without any concatenation.
    - The loss used here is categorical cross entropy.

- Summarising Results:

    - As you can clearly notice in the notebook, the validation accuracy in Model 1 jumped straight to 61% in the first epoch itself. It crosses the required threshold of 70% in the second epoch.
    - The additions that I found to be the most bearing to the customary architecture provided, were an additional flatten and dense layer with slowly increasing number of kernels in the convolution layers.
    - One of the most important steps that I feel I learned during this assignment was setting the maximum token size and output sequence length. I set mine to 12000 and 1500 respectively. Why and how? Trial and error. Moreover, I looked at percentiles after vectorizing the data for both the parameters. In addition to this, I also considered this article not only as a reference and guide throughout the procedure, but also as a sort of initialization for my hyperparameters.
    - The maximum validation accuracy and validation F1 score were 83.854% and 0.8371 respectively. The maximum training accuracy reached was 98.24%.

    - About Model 2, the validation accuracy crosses the required threshold of 10% at the third epoch.
    - It does not stop improving until the final epoch where it reaches 13.087%. The validation F1 score reaches 13.09.
    - Note that the low metrics are due to using character embeddings.