1)

The 3 mistakes in the code given are:

* The input dimensions value is 2000 i.e the size of vocabulary in data
* Output neurons should be 3 for positive, negative, neutral in the target column of dataset.
* Output layer activation function should be softmax as it works best for the multi class classification

Compressed the dataset to 2000 records.

Tokenized the data and converted into the text to matrix form.

Used the label encoder method to convert the text to digits and fit, transformed the data.

Split data into train and test which is considering the 25% as the test data.

Used the deep learning sequential model with 2 layers.

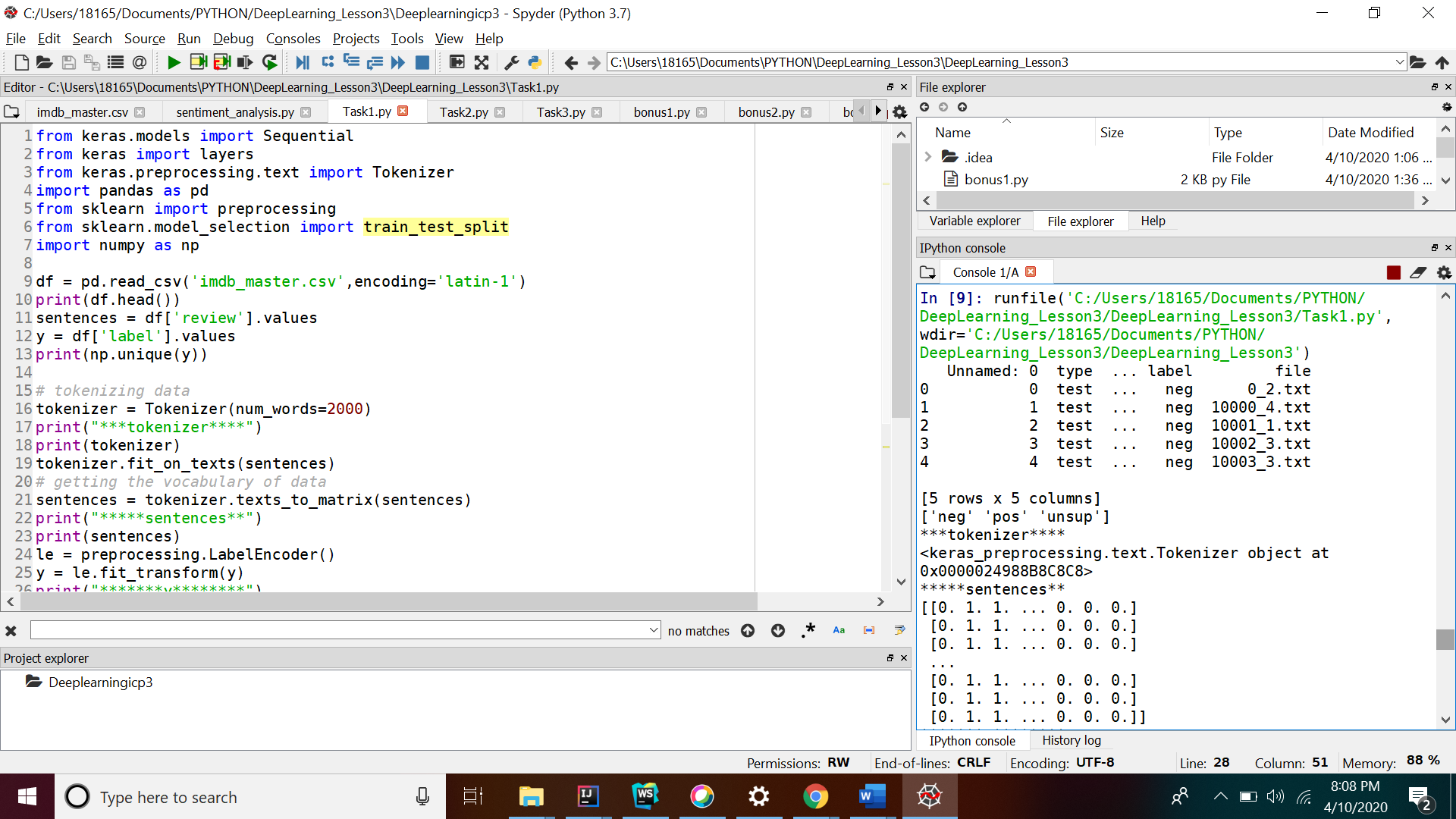
1st layer used is Embedded layer which finds out the meaning and captures the semantic relationships within the data. It follows the default word2vec algorithm which looks at the bigrams relationship among the data

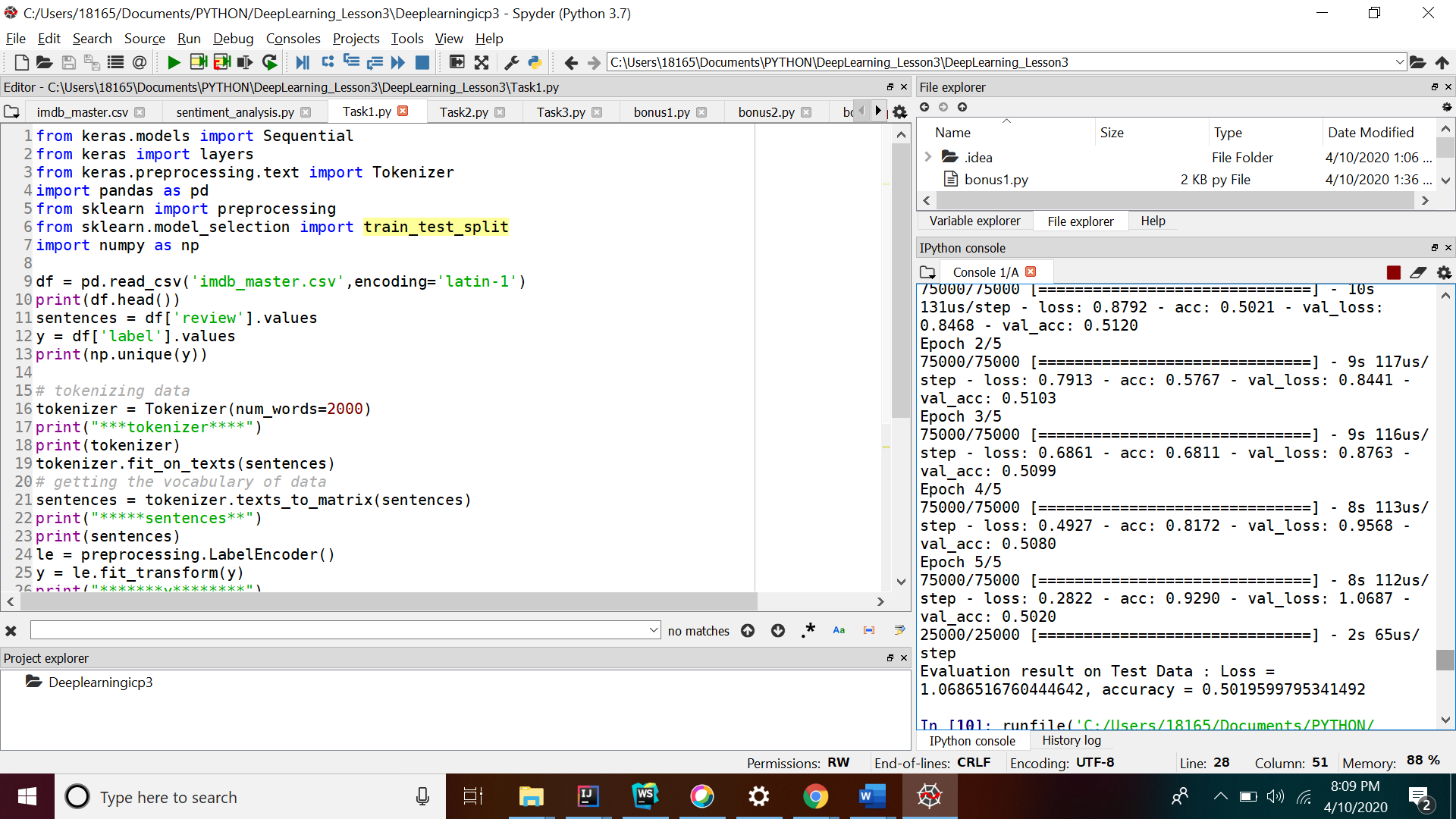
2nd layer - No of neurons is 300 and activation function is relu.

Output layer - No of neurons is 20 and activation function is softmax as the output layer

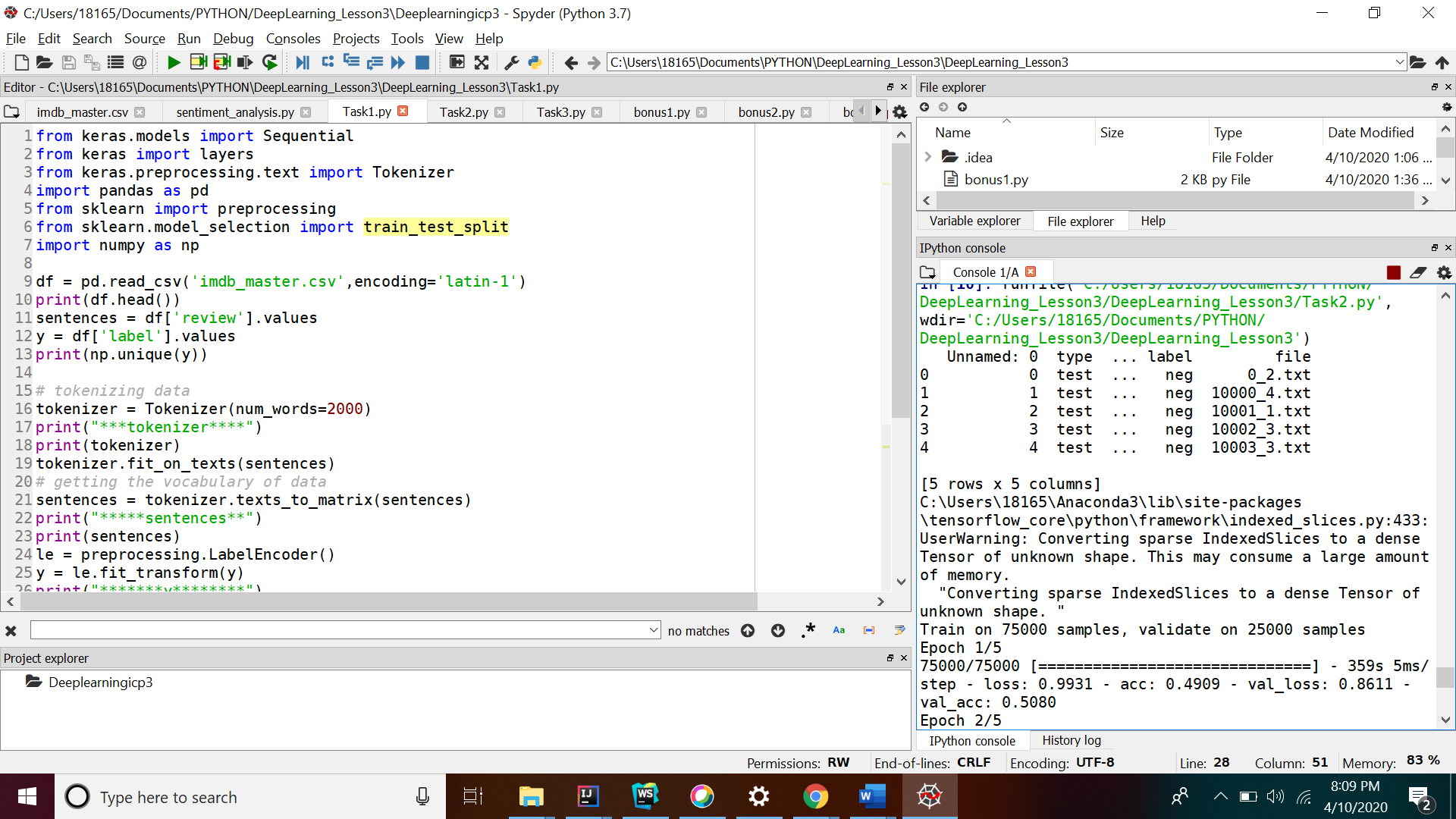
No of the epoch is 5, Batch size is 256 and loss function used is sparse\_categorical\_crossentropy

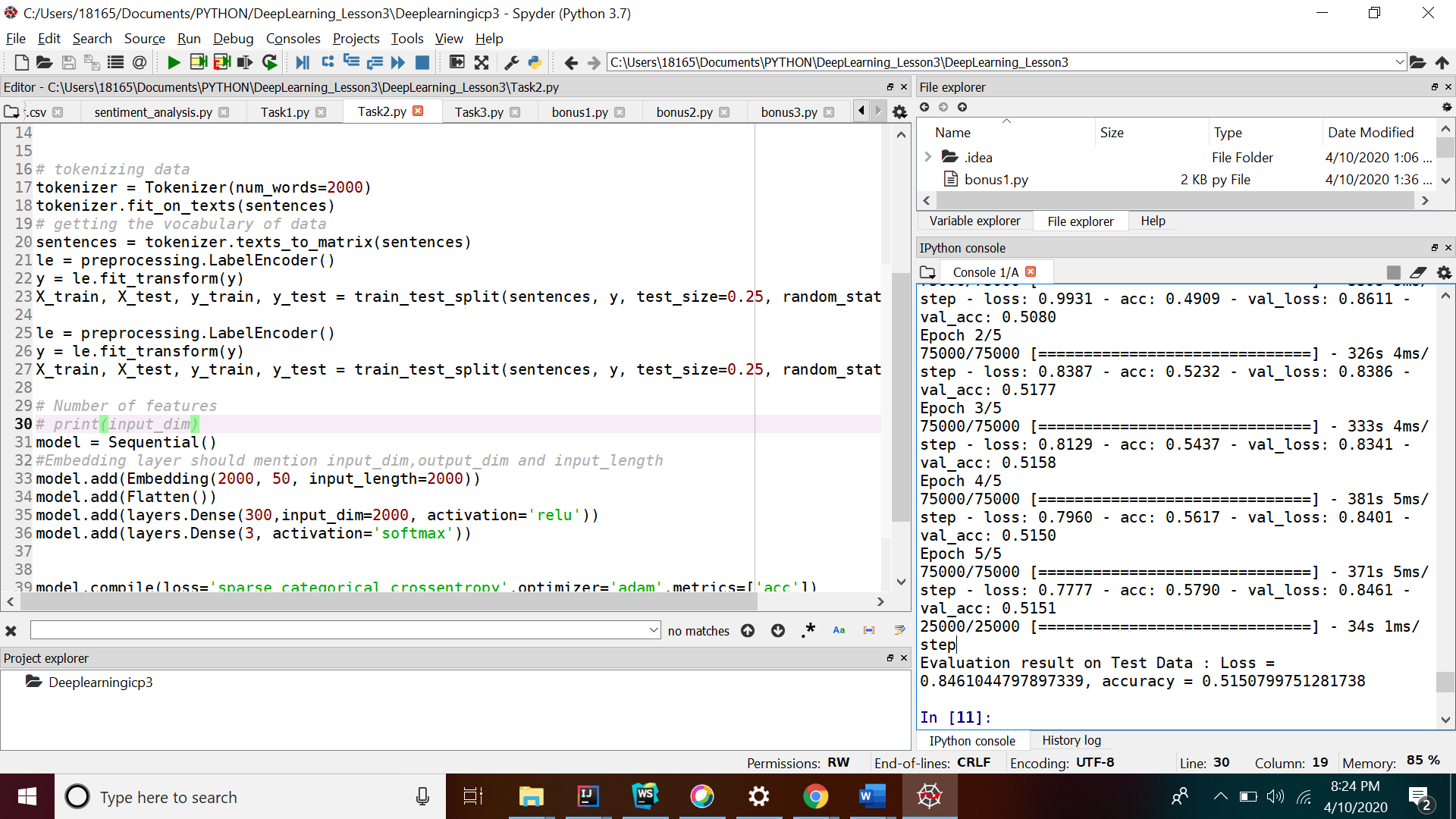
Accuracy is 50% for test data and 92% accuracy for training data



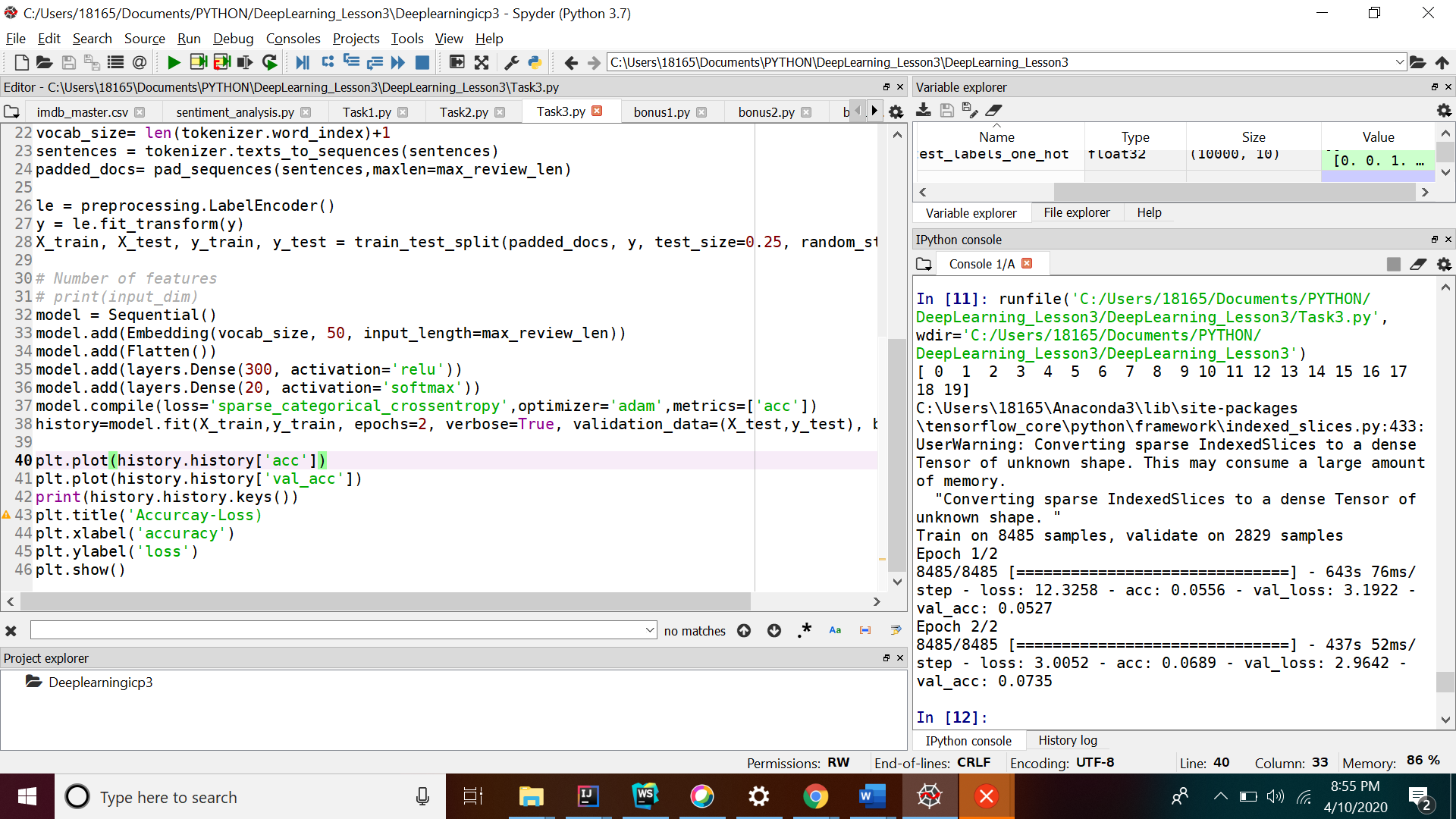


2)

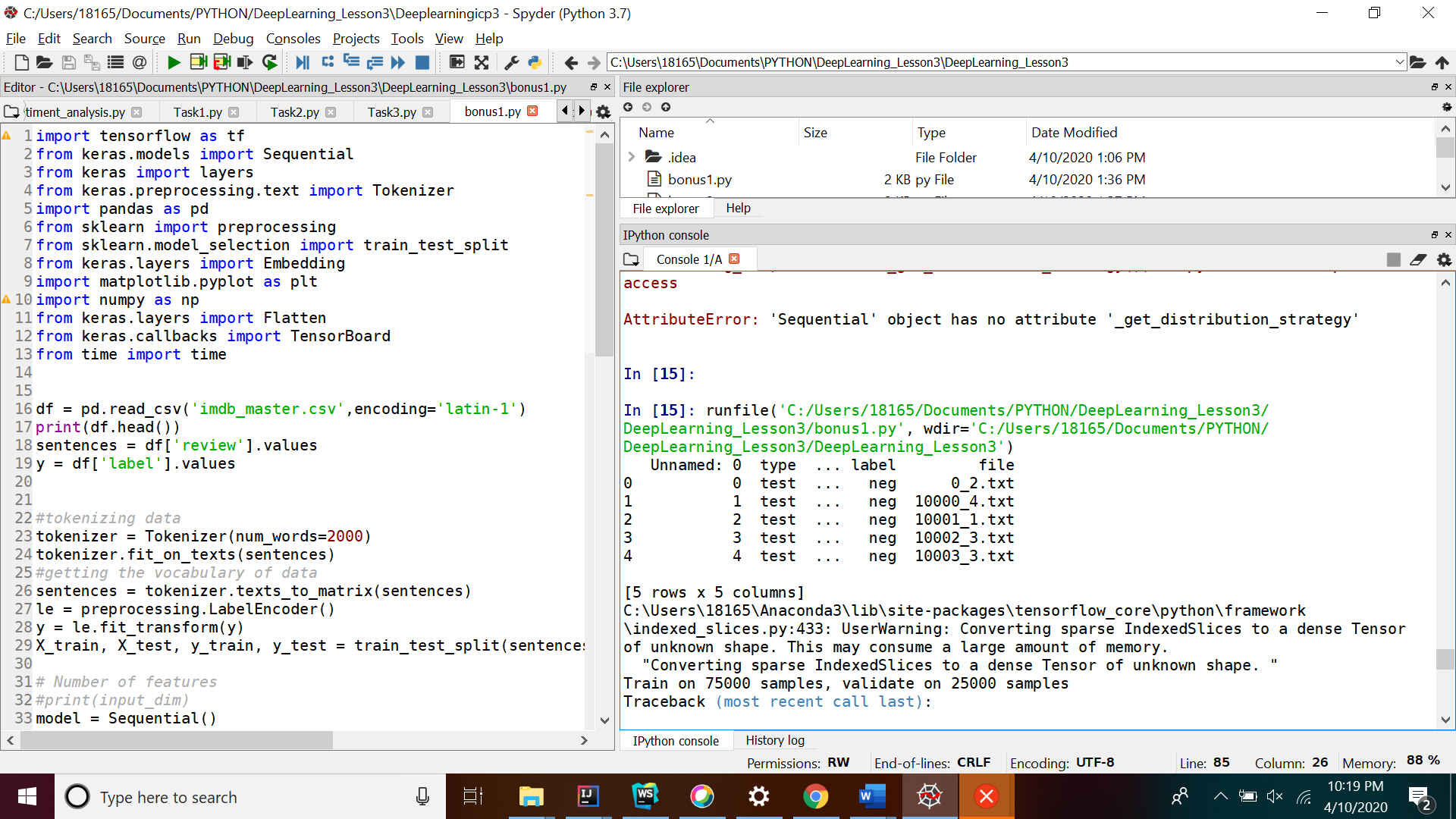


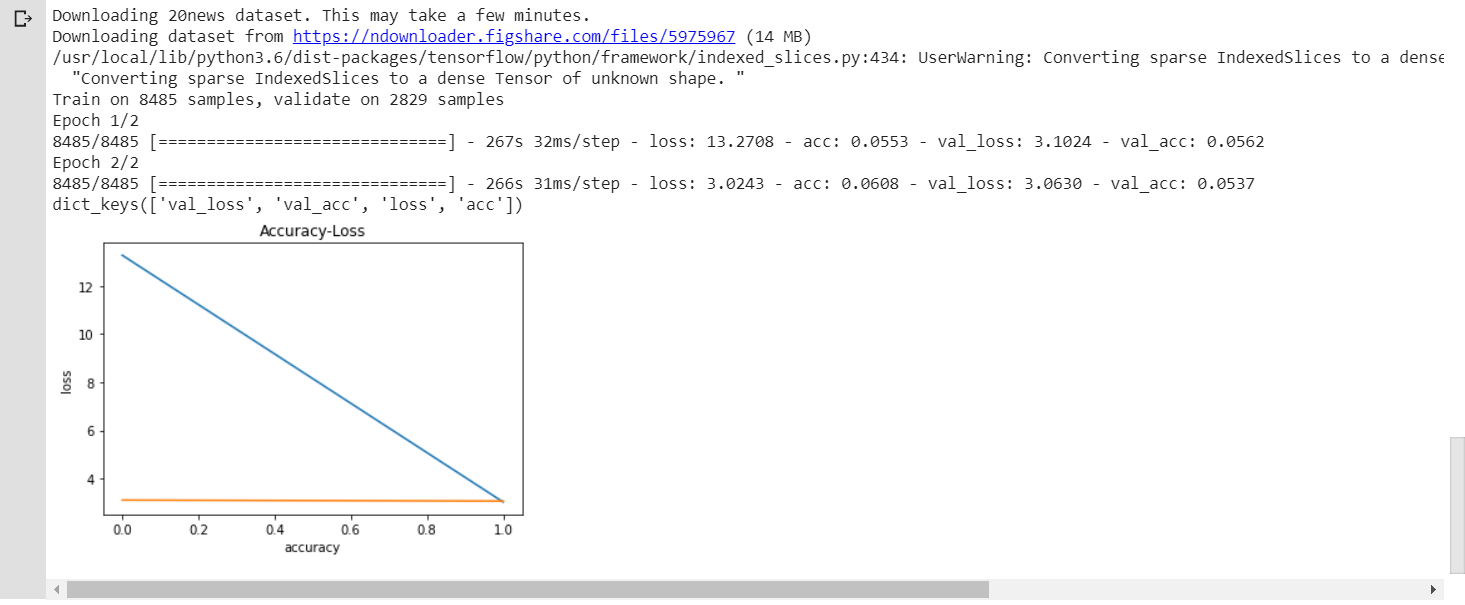


3)

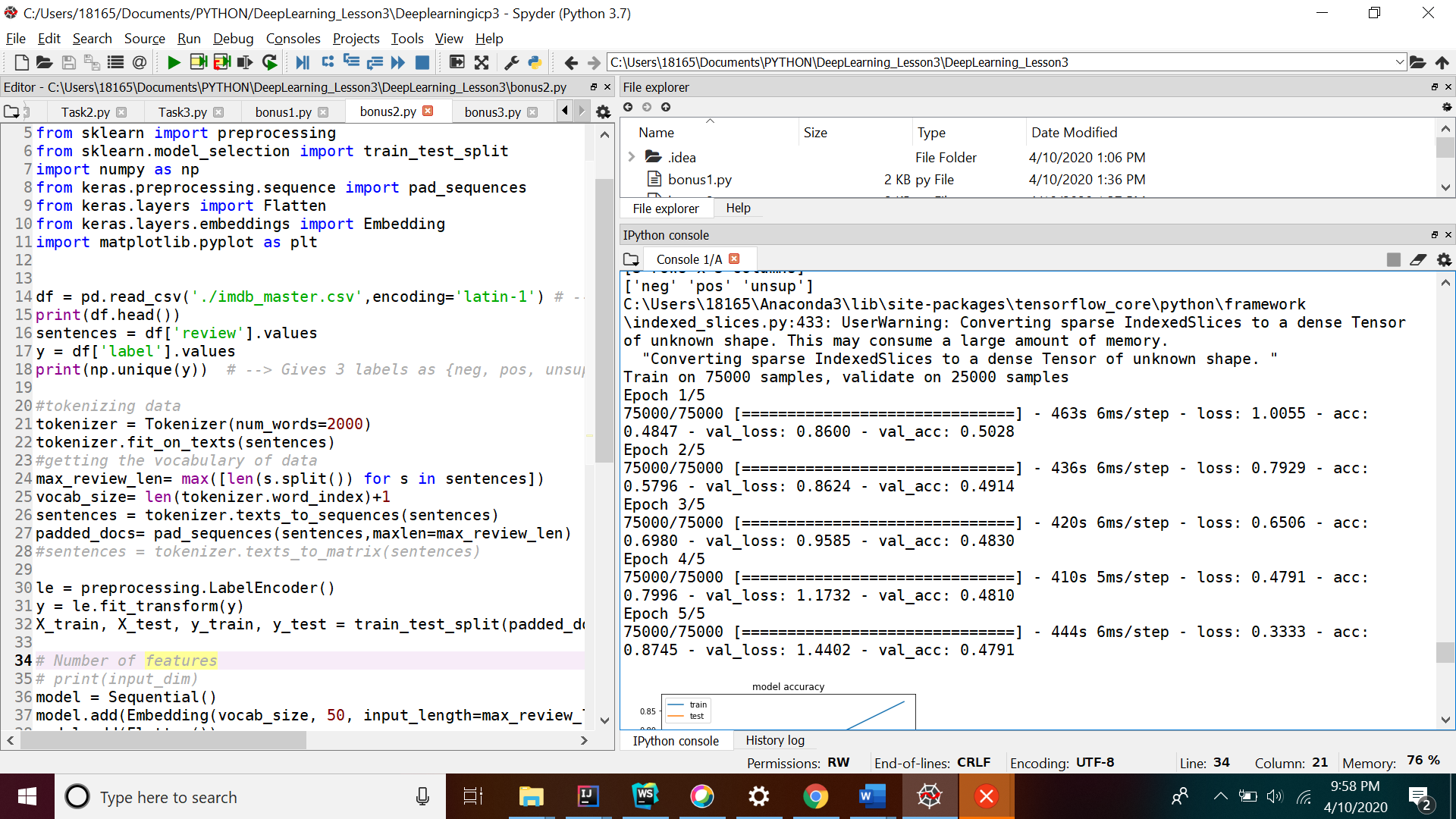


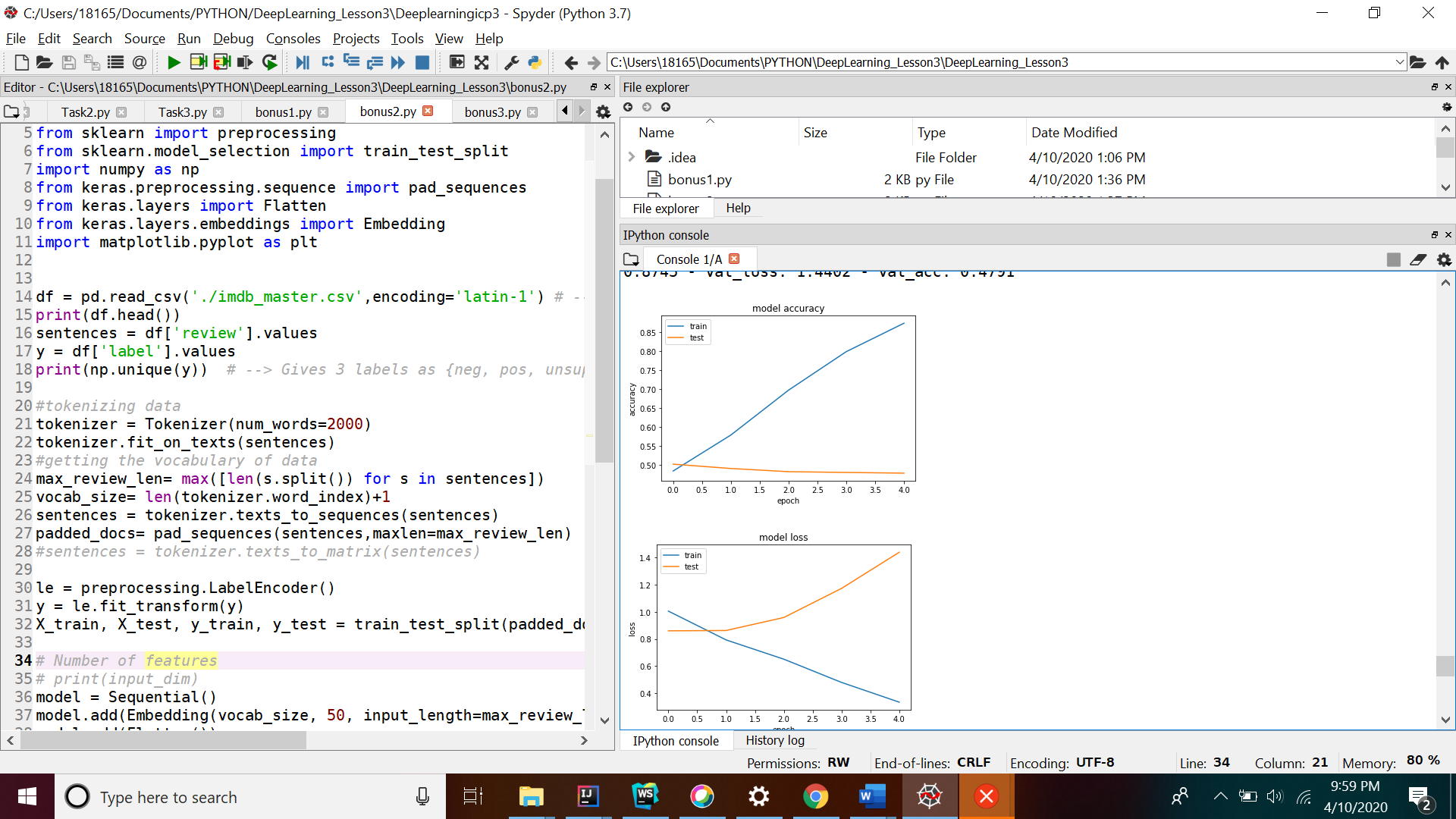
4)Bonus



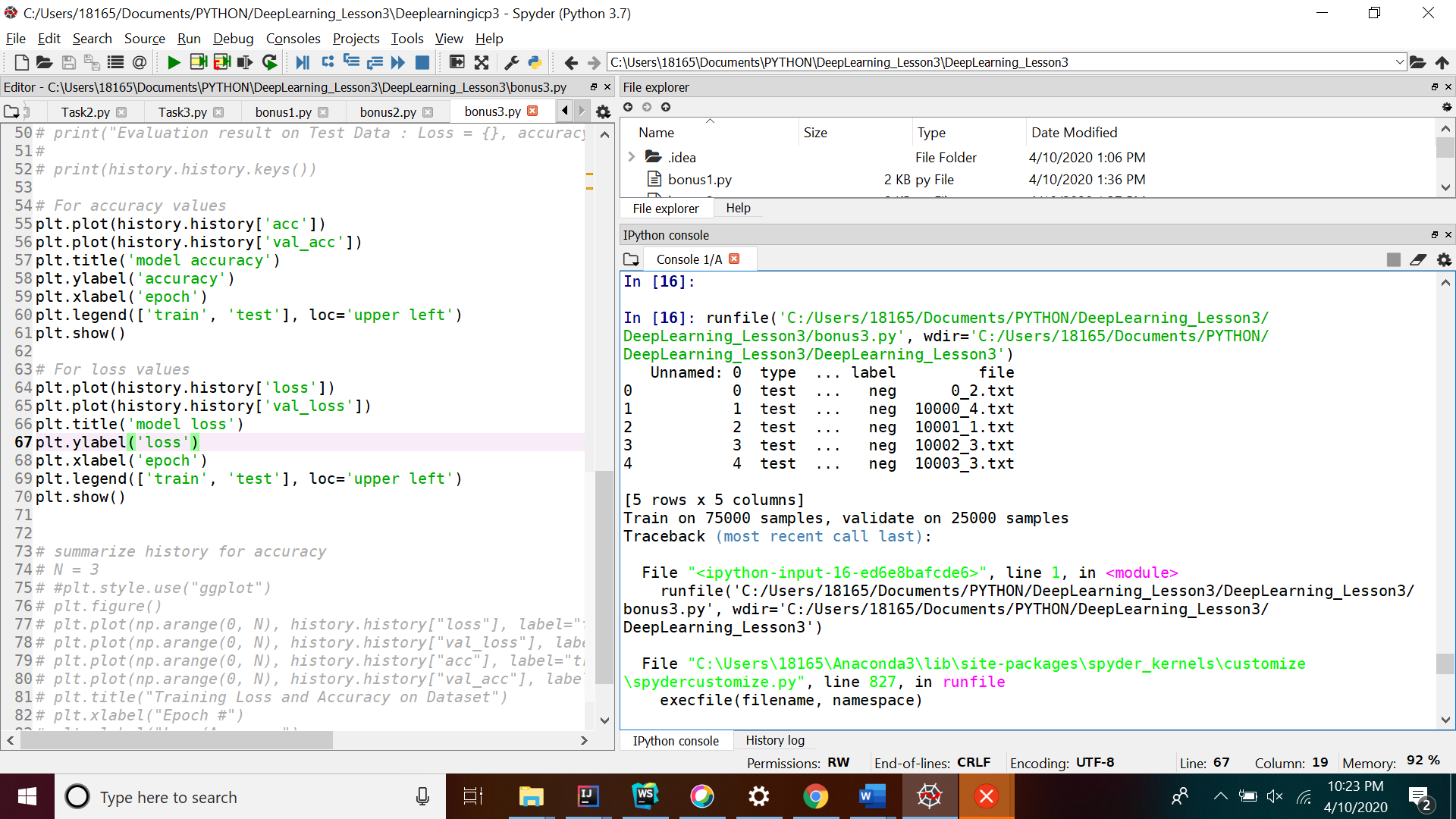


Bonus-2

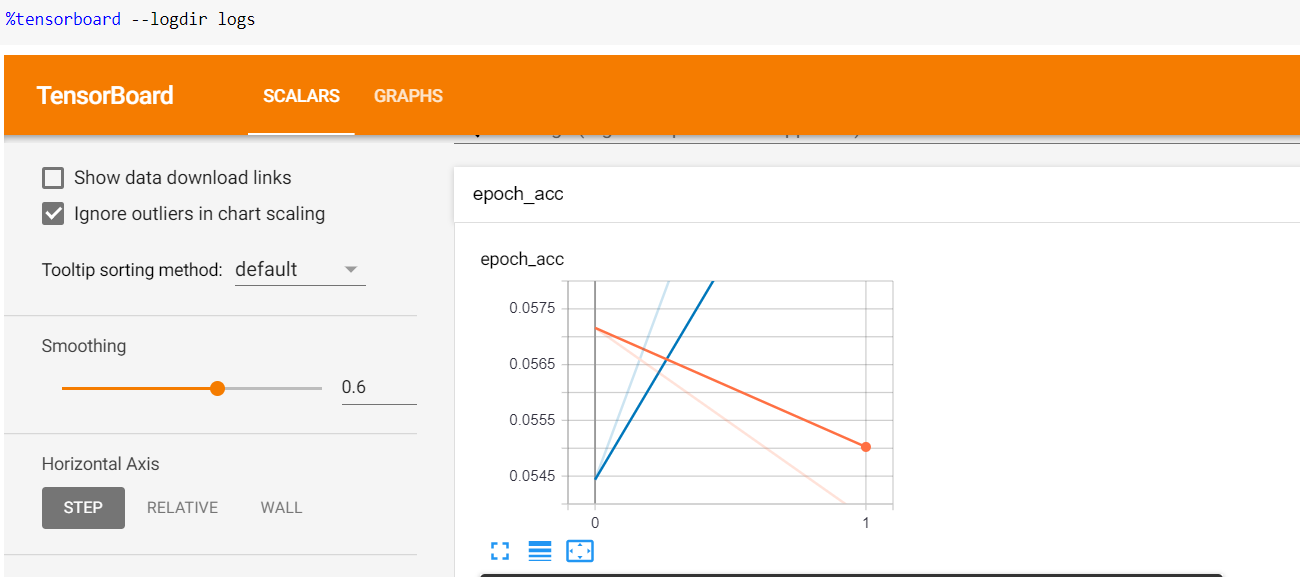




3)Bonus-3



Below is the accuracy for training and validation data.



Below is the loss for the training and validation data

