We used the Pima Indians Diabetes Database which was provided in a csv file. It has 8 features and one output - i.e. whether a person has diabetes or not

The task was to build two Ml models for prediction of the same based on the features.

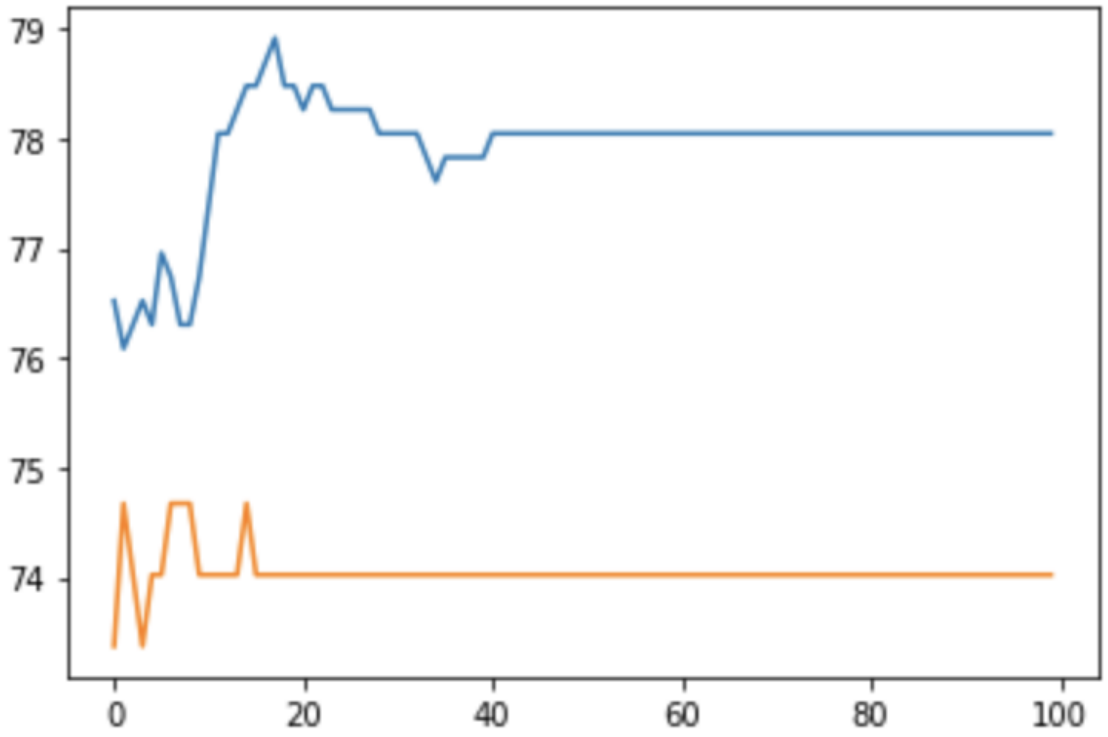
Logistic Regression

I got a test set accuracy of 78.57%

Initial learning rate was set to 0.101
Decay was set to 0.1 after 50 steps
Number of epochs were set to 100.

As can be seen in the below chart, I could have stopped the training at early as 50 steps as there were no changes to the model after that

Attaching a graph which shows the training and validation accuracy while training



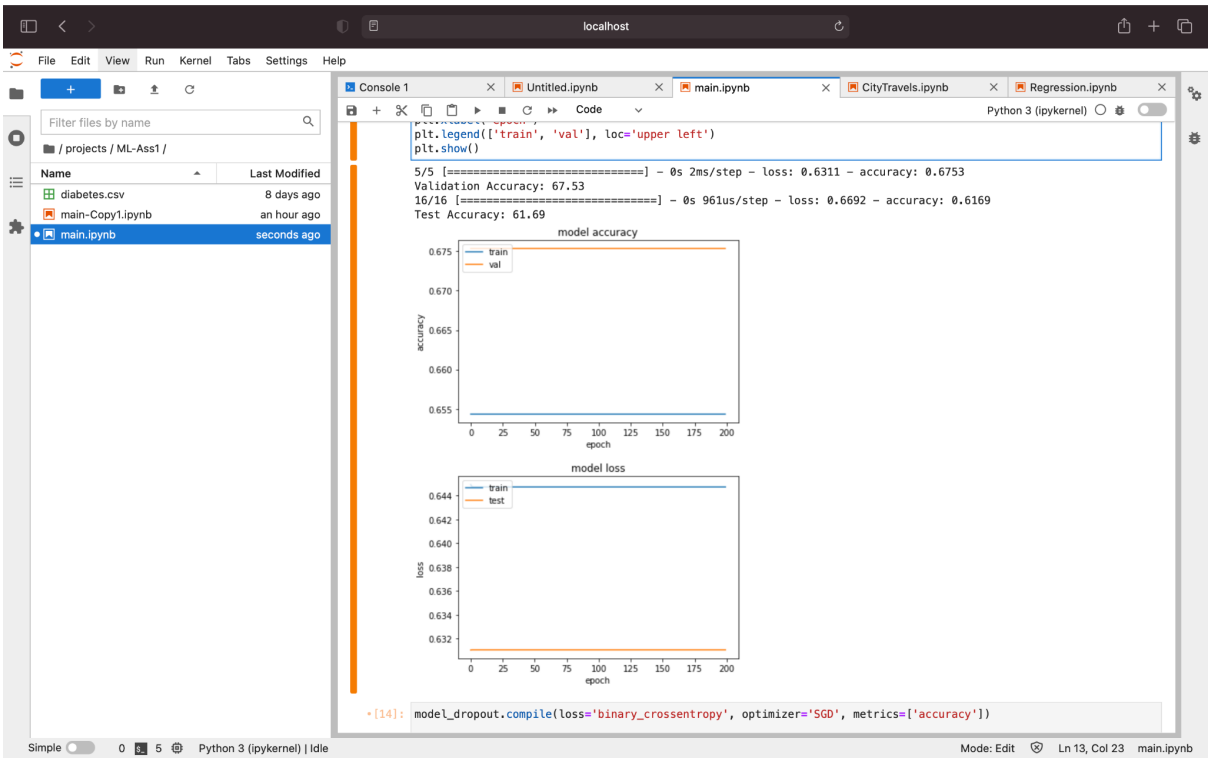
Blue line shows the training accuracy, and orange shows validation accuracy

Neural Network

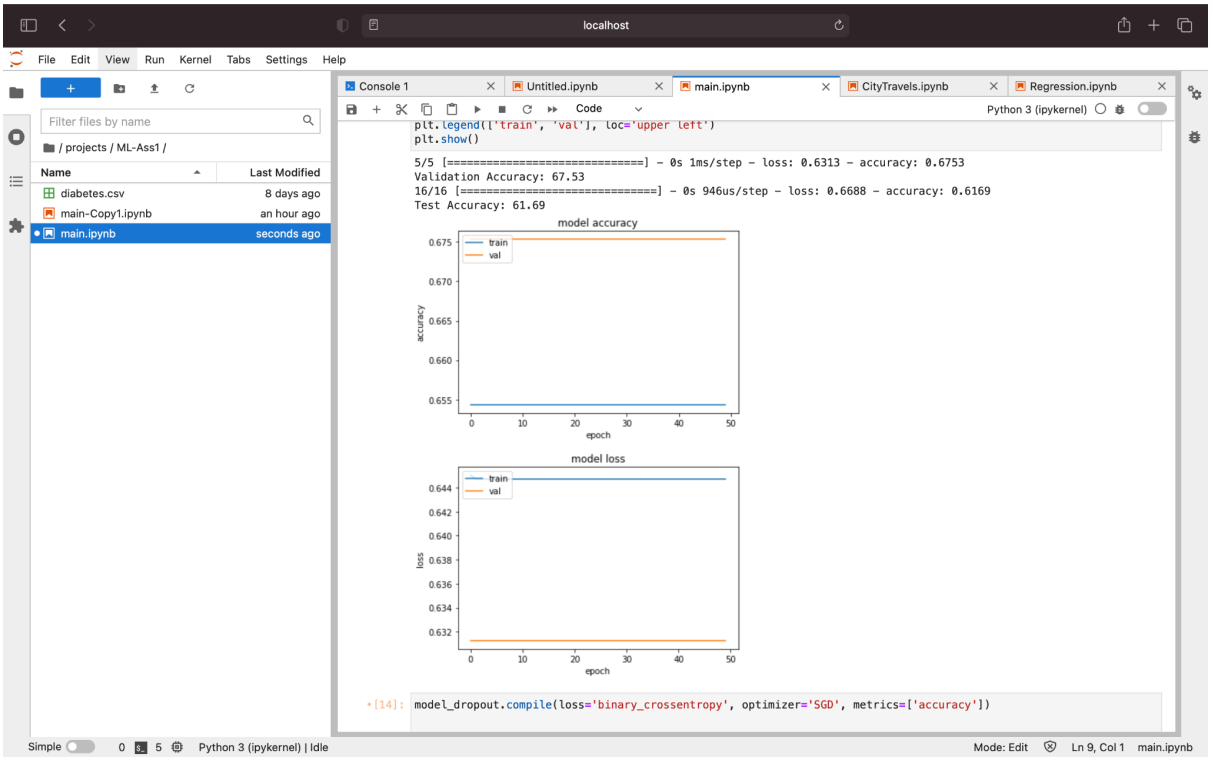
In Neural Network, I tried four different models. All of them have 3 hidden layers with 10, 8 , and 4 neurons respectively.

1) **Without regularization**

This preformed the poorest of all. **I tried learning_rate = 0.5*10^-2, with decay_steps = 50, decay_rate=0.7 and number of steps = 200**. Attaching the graph having train vs val loss and accuracy

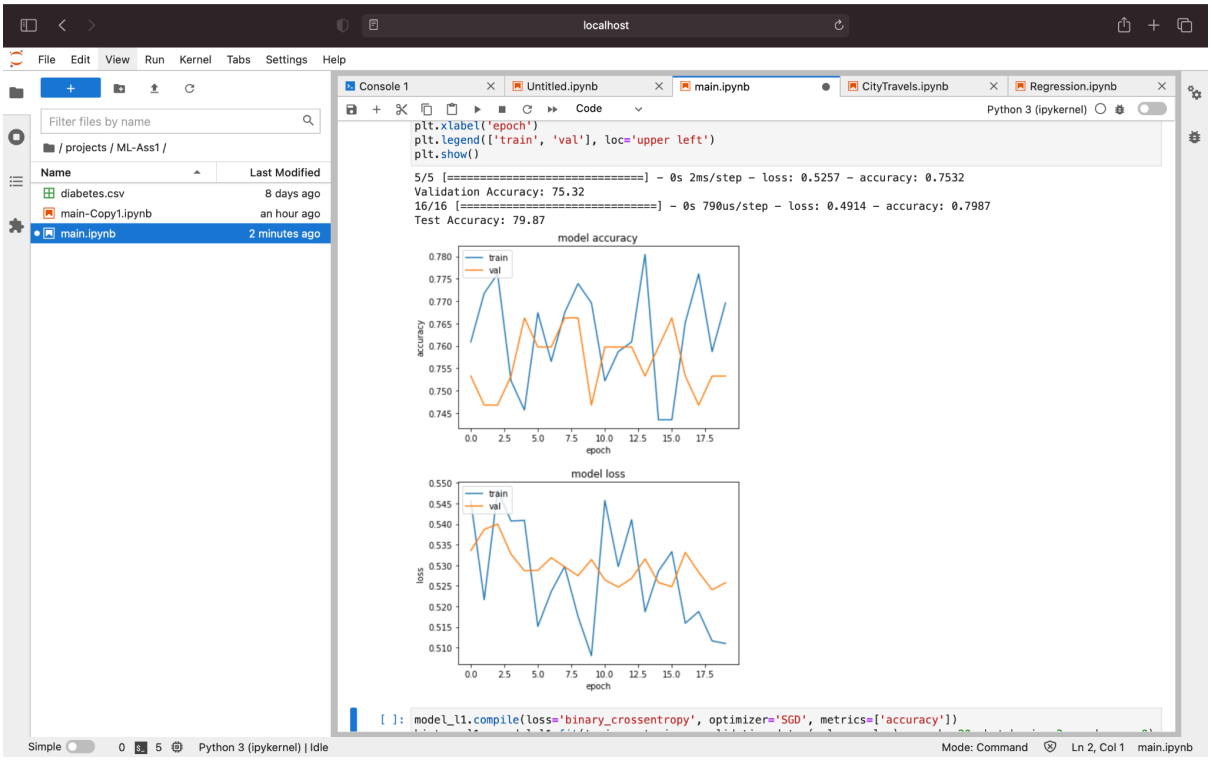


I tried changing decay_steps = 10, decay_rate = 0.9 and number of steps to 50. But the graph remained similar.



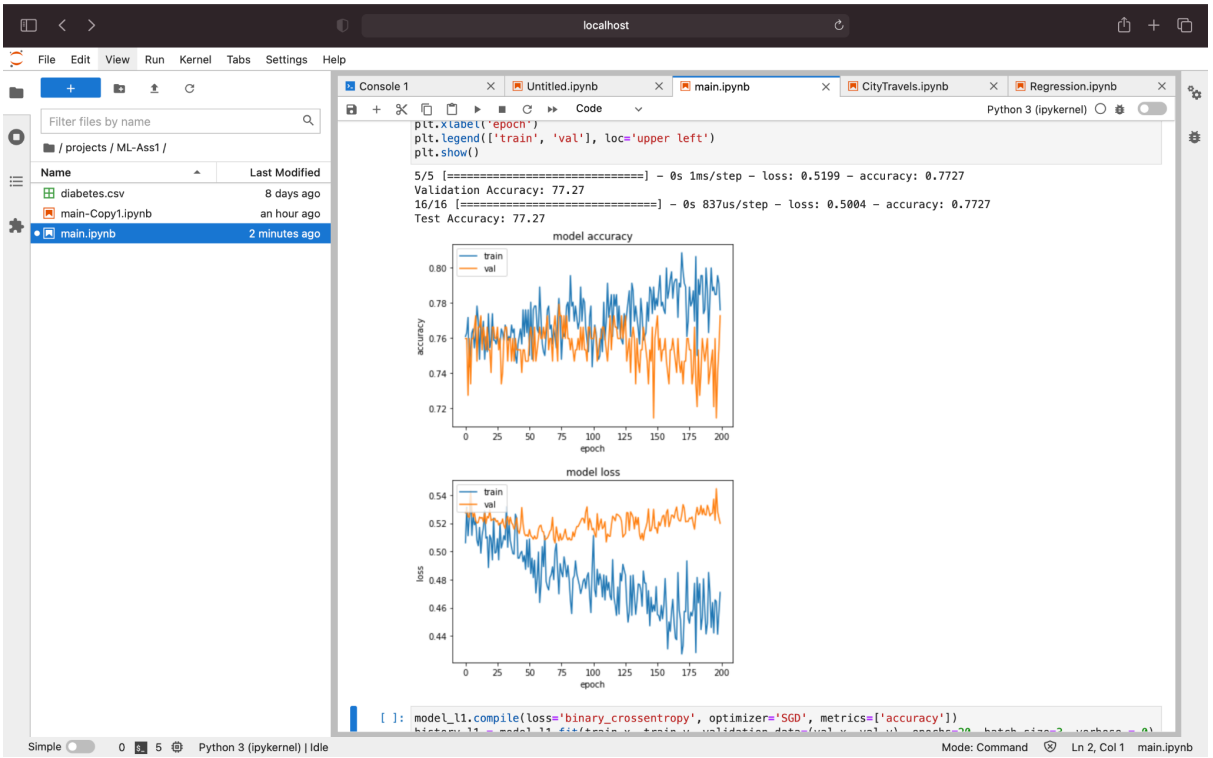
2) With only Dropout - 0.1 after every hidden layer

This performed the best, with val and test accuracies of 75.32 and 79.87 respectively.

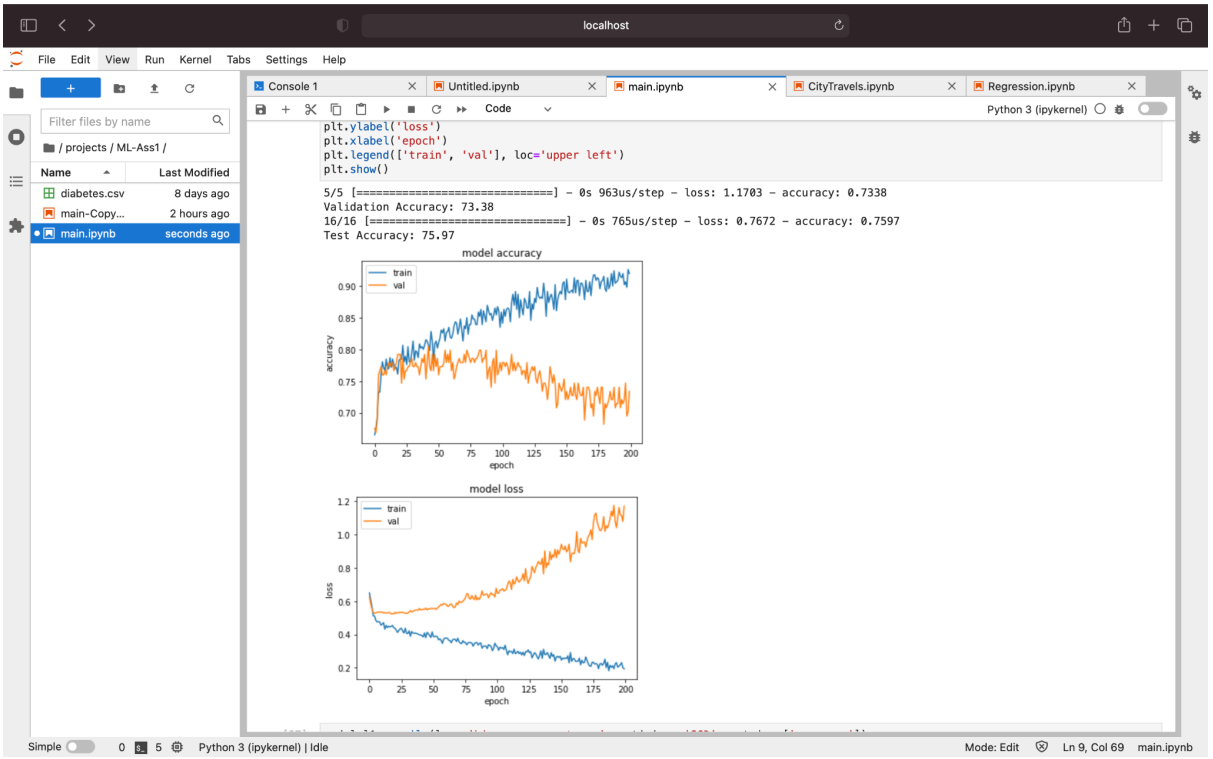


It uses SGD and run on just 20 epochs. Seeing the loss graph, it looks like the model should benefit from more training. Hence I ran it on 200 epochs, and the **val and test accuracies both became 77.27.**

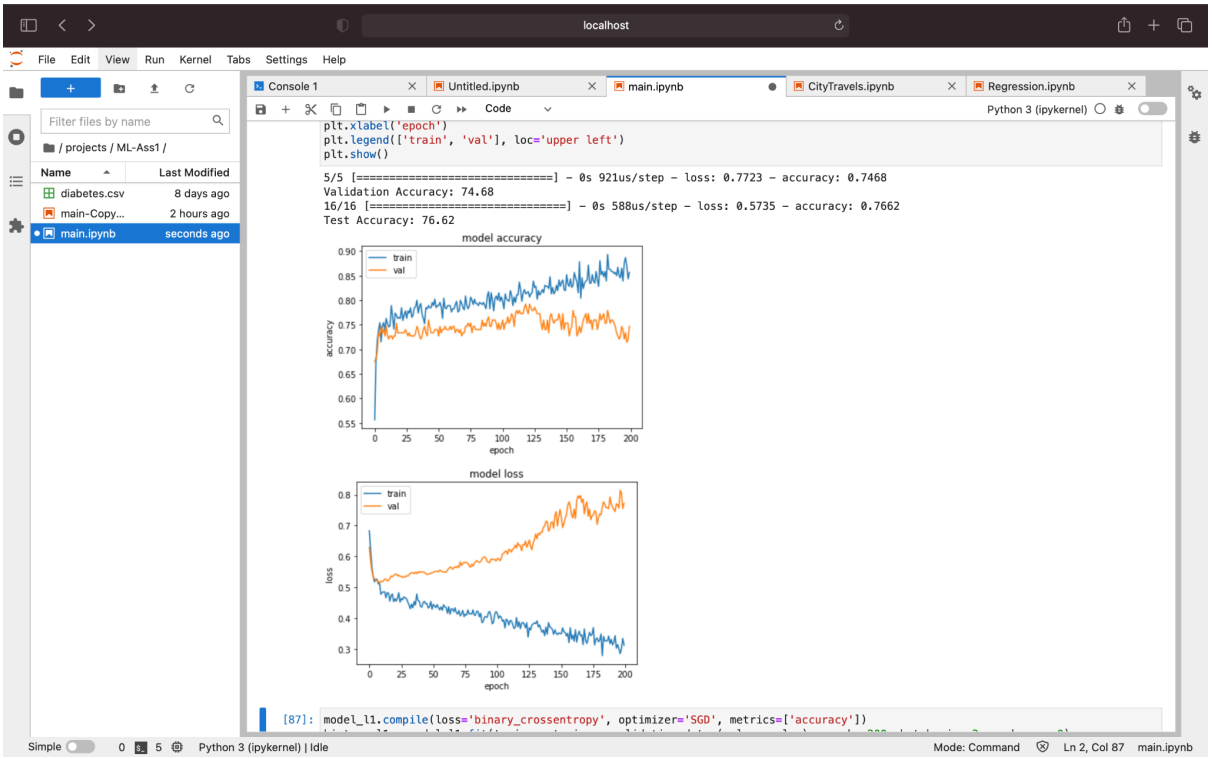
Attaching the graph below



After this, I tried changing the number of neurons to 80,20,and 10 respectively keeping other variables same. Attaching the results below

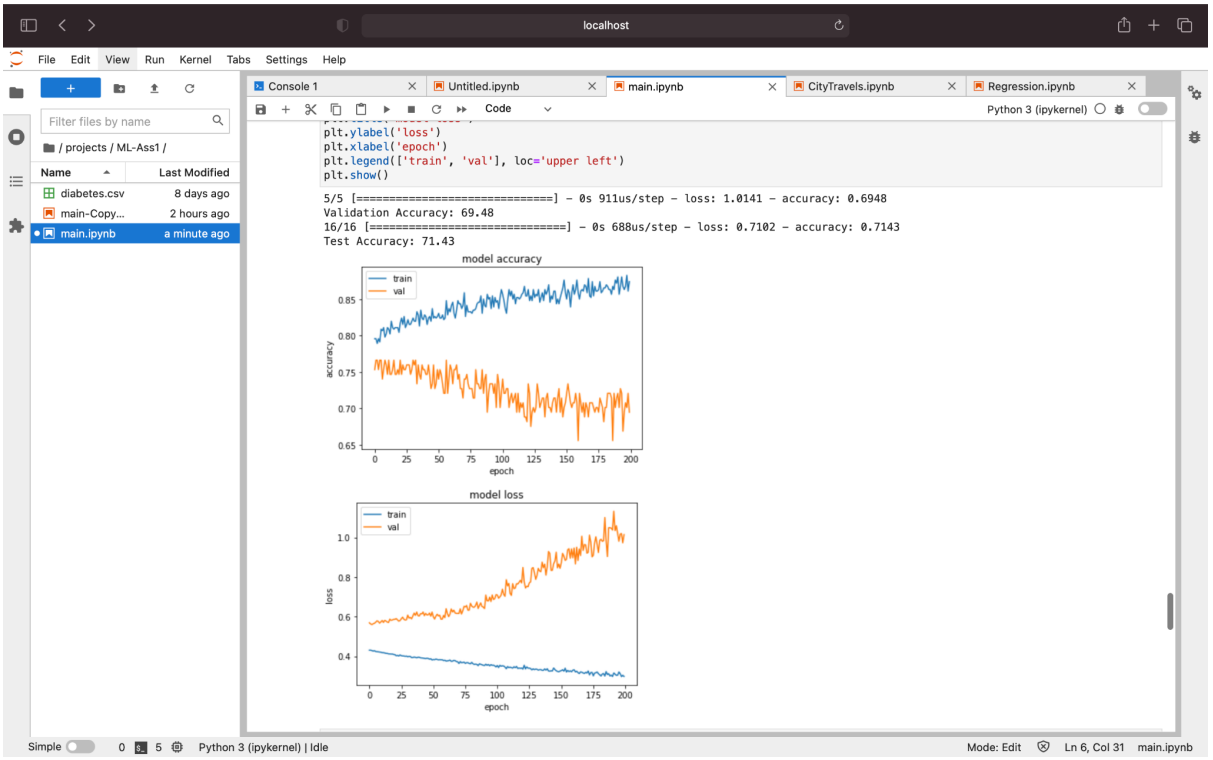


The accuracies decreased, probably because it is not so complex model and maybe if I increase the dropout, the results might improve. Trying with dropout - 0.2 after each layer

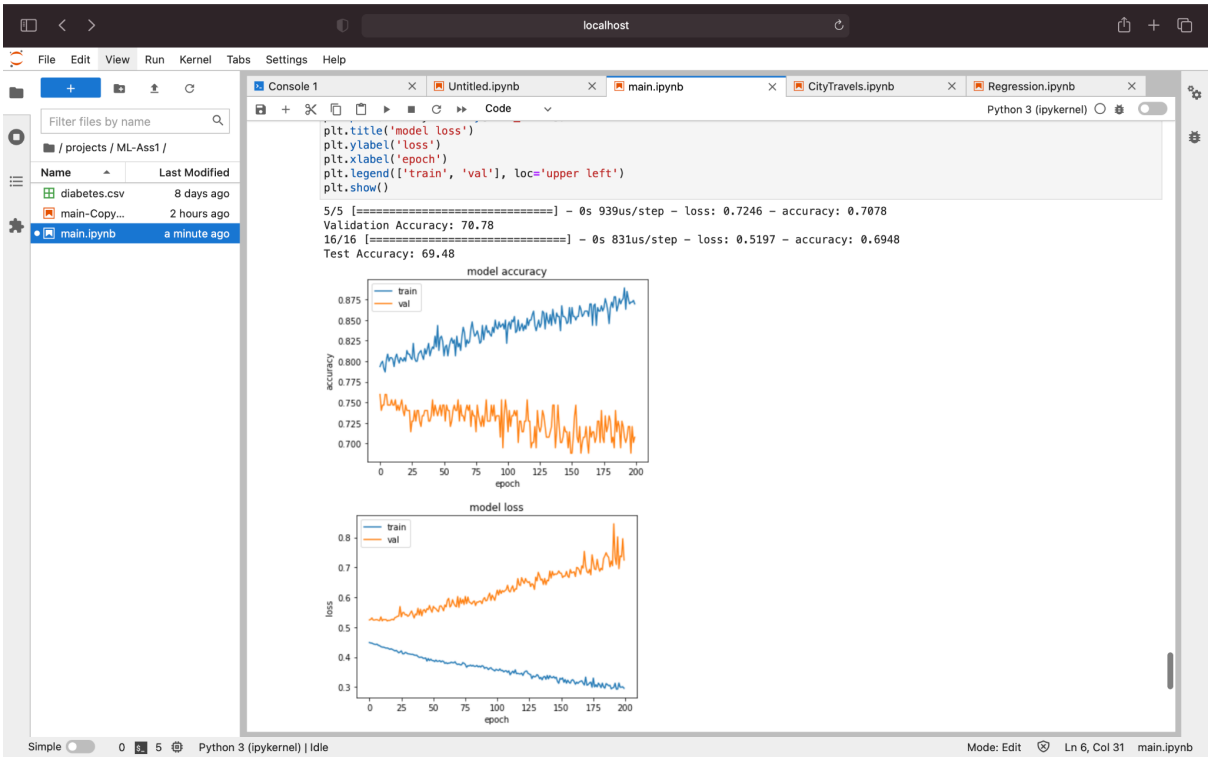


Val accuracy increased but they are not as good as our original dropout model

3) With only L1 - 10^-4 at every hidden layer



4) With only L2 - 10^-4 at every hidden layer



Dropout vs L1 vs L2

Accuracies