**Neural network**

In this project, we took the titanic dataset, and we want to predict who survived and who didn’t. We have to mention that our dataset is small.

We started by cleaning the data, we dropped the rows with NAN values and we dropped columns that can’t help us predict if the person survived or not like: ‘Fare’, ‘Cabin’ and ‘Ticket’, which depends on ‘Pclass’ column and ‘PassengerId’, ‘Name’ and ‘Embarked’ which were not helpful for our model. We split the ‘Pclass’ column into three different columns and we named them ‘Pclass\_1’, ‘Pclass\_2’, ‘Pclass\_3’. Then we changed the 'sex' column into 0,1 (0=female, 1=male) and last we normalized the ‘age’ column.

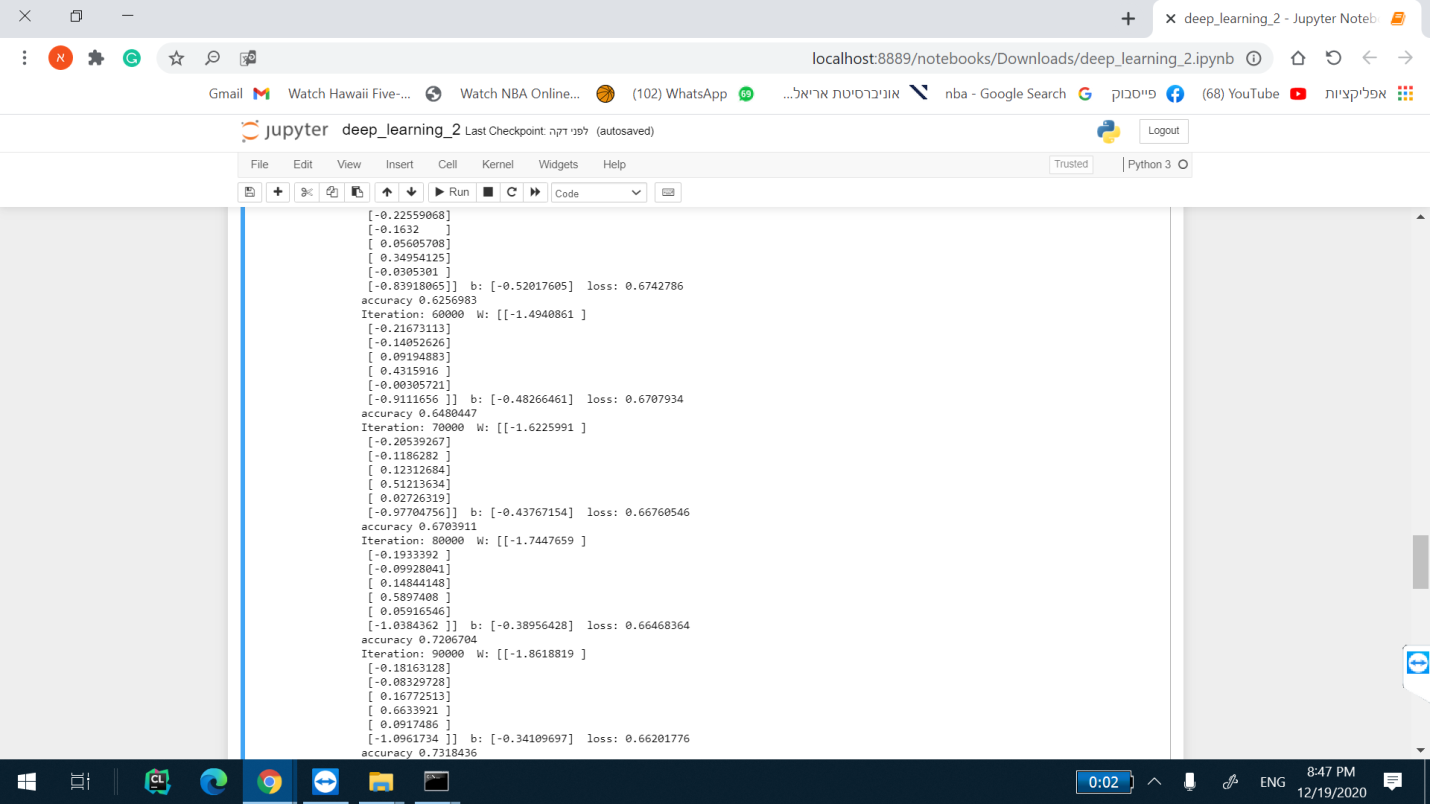
After we cleaned the data we set up the x and y labels, and then we split them into x\_train, x\_test, y\_train, y\_test (test = 0.25, train = 0.75).

We chose to use the logistic regression model on our data and we did 100000 iterations and 0.001 learning rate after we tried a few times with different numbers. In the end, we could see the weight, the bias, the loss and the accuracy of our prediction.

Then we wanted to improve the accuracy of our prediction so we added one hidden layer with 3 nodes and 1000000 iterations and 0.001 learning rate after we tried a few times with different numbers. We chose to train this model more iterations since it is a more complicated model.

As we can see using hidden layer after 300000 there is overfitting, so it will be best to stop before. As we can see, compare to the logistic regression model there are improvements in our predictions.

Screenshot of the results whith logistic regression



Screenshot of the results while using hidden layer

