

In this project, I implemented a multilayer perceptron with 3 layers(1 hidden layer). The hidden layer has 5 units. The model can classify input into 3 classes with 2 features with above 70 percent accuracy. 400 data points are labeled and available.

I splitted the data set into 2 parts : Training set and test set. I used 80 20 ratio when dividing. I found this ratio from a question at stackoverflow: <https://stackoverflow.com/questions/13610074/is-there-a-rule-of-thumb-for-how-to-divide-a-dataset-into-training-and-validation>

I did not used validation set because data is very few and it is not mentioned in the project description.

I used mean square error function. I used xavier weights initialization. I used gradient descent as optimizer.

Hidden layer and input layer have extra 1 unit that is always one which is for bias.