

ITF31519 - Assignment 2

Tobias Hallingstad

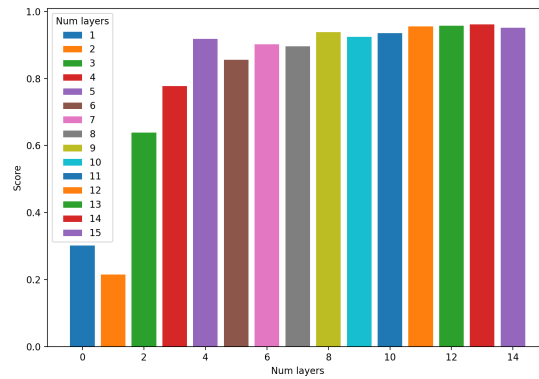
October 13, 2021

1 Default values

I am basing the `MLPClassifier` on some default values. These values are used if no other value is specified: `randomState = 1`, `maxIter = 300`, `nLayers = (100,)`

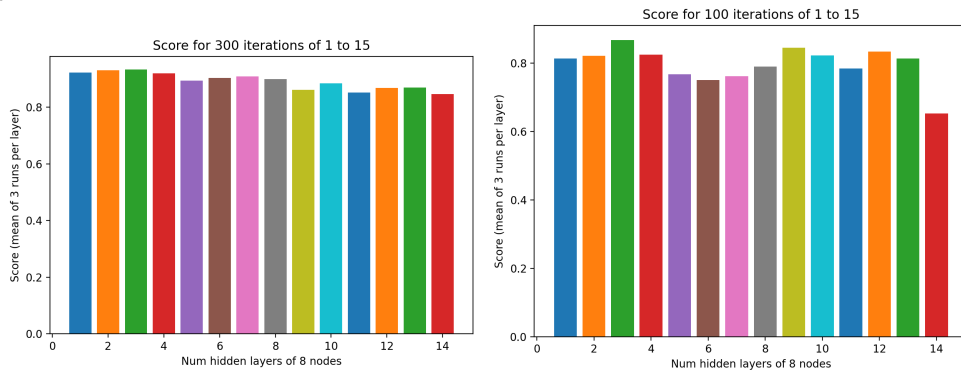
2 Observations

Number of nodes I did some testing on what effect the amount of nodes have on the model. I wrote a function (`plotMulti2()`) to run multiple iterations of the model (with different number of nodes), then plot the result in a bar diagram. Here I run 15 different models, where I increase the amount of nodes in the first layer.



Here I can see that after 8 nodes there is little change in the score.

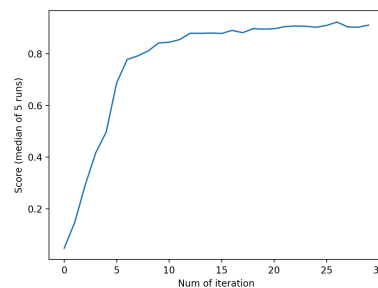
Number of hidden layers After testing how many nodes in 1 layer looks good. I proceeded to test how many hidden layers I can use with 8 nodes in. For this I used the `plotMulti3()`. Here I played around with the `iterations` variable to find some interesting data. I here use 150 to see if that changes the score.



The graph show that when this model gets more layers (here of just 8 nodes) we see that the score gets worse. I would think this is because the model is trying to learn more then what it should, so it guesses wrong for data that is not in the test data.

Accuracy For calculating the accracy I first just calculated the accracy using the `clf.score(testX, testY)`. This is calculated in the `runMLPC()`. After seeing the data I wrote the `run()` too allow me to take the mean of `X` amount of iterations. For this dataset I found it hard to finde some parameters that made the score bad. I think this can be seen in the tables shown.

Using the values from the previus tests, 8 nodes in a layer with 6 layres. I show that when the iteratins are low the model is bad, but the model rapidly changes to a "stable" model after it goes true some iterations. Here I am using the `plotThing1(30, 0, 0, iterationIncreas=10)`. This means that it loops over 30 models and plots their score. Each loop the model iteration is increased by 10 and the first model is iterating 1 time. The score is the mean of 5 runs with the same parameters.



This graf shows that when iterating 150 times seems like a good choice for a good score.

Conclution It seems like runding this model with 8 nodes in 4 hidden layers and training the model for 150 iterations gives a good score of 87.78%. Increasing any of the parameters will give a better score. I would recomend to change the "number of nodes" and number of iterations.