Machine Learning Project

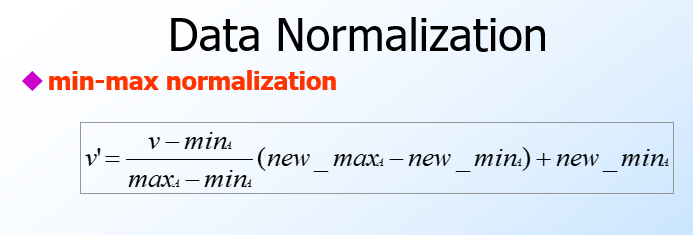
* Data Preprocessing

There are a lot of non-numeric values that need to be converted to integer form, but the problem arises that we might give a value of any given attribute a higher number which may have lower effect on the possibility of the loaning to be fraudulent.

Solution: We separate out the fraudulent example that for each value of each attribute calculates it probability and the one with the highest get the max number while the rest are given in the decreasing order.

Since neural networks work better on an input between zero and 1 we convert the data

By using the formula



Since we don’t want anything to be zero for more any attribute that has more than 2 value we use new\_min as 0.1 and for binary as 0.5.

* Training/Testing the neural network

So in the original data set the fraud loans are labelled as 2 and non fraud as 1 so to make it binary 0/1 we just (target-1) in the formulae.



Which makes it o\*(1-o)\*((t-1)-o).

We train the neural network for 10k epoch at a high learning rate of 0.5.After testing the data is seems the step function should have threshold value of 0.4. smaller than 0.4 is fraudulent and greater is non fraudulent.

README: keep learning rate high >0.4