

CENG 414
Introduction to Data Mining
Fall 2019-2020
THE 2

Question 4

1- How many hidden layers and hidden nodes are created?

There is a 1 hidden layer and there are 11 hidden nodes.

2- Did Weka normalize the attributes? What is the effect of normalizing the attributes?

Yes, Weka normalized the attributes. In order to make all inputs in a comparable range, normalisation is required. If normalization process is not applied to inputs, importance of inputs are not equally distributed.

3- What is the benefit of splitting the dataset as training set and test set? Why don't we just train our model with whole data?

If we train our model with whole data, our model memorizes the training data. However, when the model wants to be tested, the same inputs given the program as well. Our model will be work with %100 correctly in that situation. But this is not the case. If different inputs are given to program, then model may be give false output.

4- Which halting strategy did MLP use?

WEKA uses the validation threshold parameter for multilayer perceptron, which is 20 in default. This validation threshold parameter is the number of consecutive increases of error allowed.

5- What is the detailed accuracy table by class of the run?

TP Rate -> It means true positives. Rate of classified elements correctly as that class.

FP Rate -> It means false positives. Rate of classified elements incorrectly as that class.

Precision -> (given class instances correctly) / (total instances classified as that class)

Recall -> Same as TP rate. (Instances classified as a given class) / (actual total in that class)

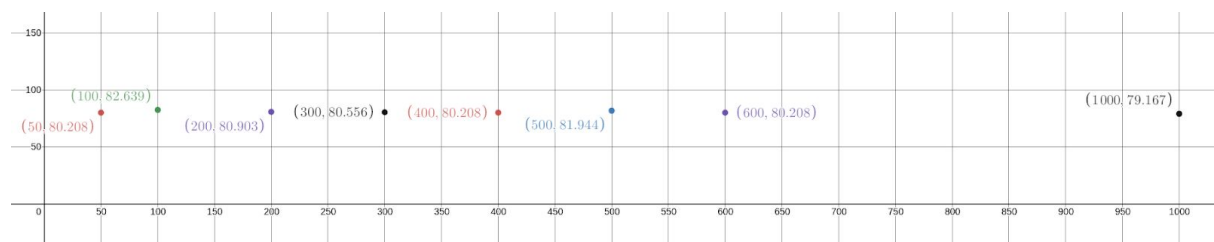
F-Measure -> Calculated by using recall and precision. It's value is $(2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$.

MCC -> measure of the quality of binary (two-class) classifications.

ROC Area -> Receiver Operating Characteristics area. General information about how the classifiers are performing in general.

PRC Area -> Precision Recall area. Shows the relationship between precision and sensitivity.

6- Plot the training time-test accuracy plot.



The x axis represents training time and y axis represent accuracy percentage.

I tried training time values for 50,100,200,300,400,500,600 and 1000.

Accuracy results are 80.2083, 82.6389, 80.9028, 80.5556, 80.2083, 81.9444, 80.2083 and 79.1667 respectively.

7- Interpret the accuracy plot: What is the relation between accuracy and training time (epoch count)? What may cause this situation?

When training time is increased, accuracy will not be increased for all the time. The reason for this is known as “overfitting” in machine learning jargon. When the model is trained more than optimal time, model is starting to memorize some exact data, not the expected common features detection.

Alper KOCAMAN - 2169589