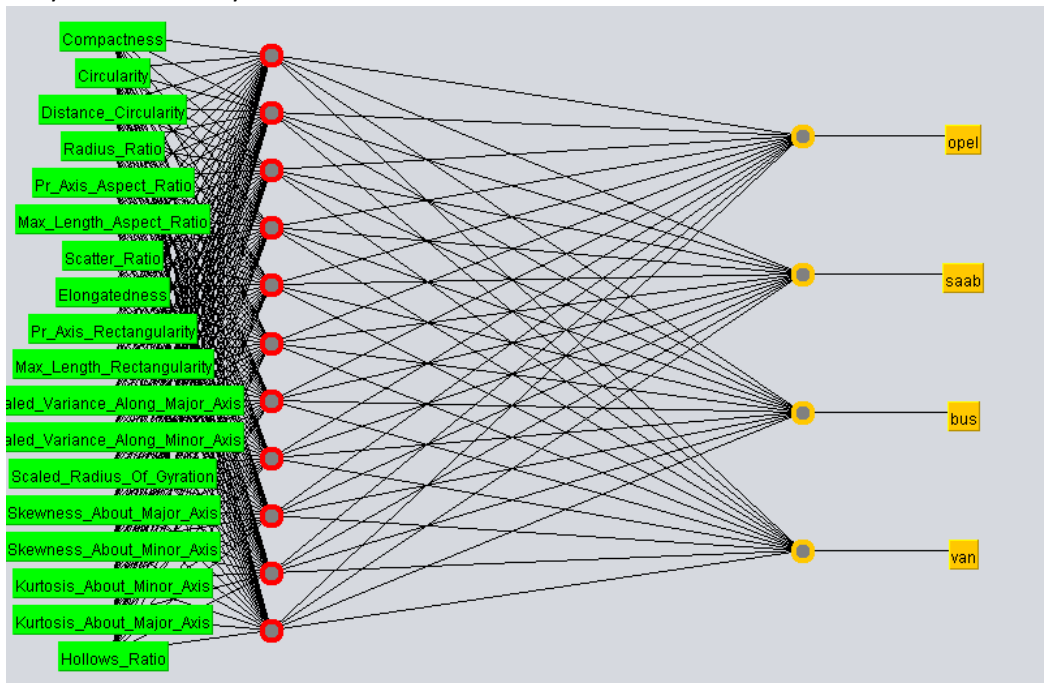


- 1- Only one hidden layer is created with 11 hidden nodes.



- 2- Weka didn't normalize the attributes. Maximum and minimum values of attributes are not between 0 and 1.

Normalizing data rescales values of attributes into an interval between 0 and 1. We can use normalization when we need to equalize range of attributes with different ranges. Normalizing effects the accuracy positively in classification.

- 3- If we don't split data into training and test data, all of data becomes training data. This means, our inference works very well on given data. However, we can't know how good is our inference on untested data. As a result, we need to split data into training and test data.

- 4- 'Early stopping' method.

- 5- TP Rate: rate of true positives (instances correctly classified as a given class)

FP Rate: rate of false positives (instances falsely classified as a given class)

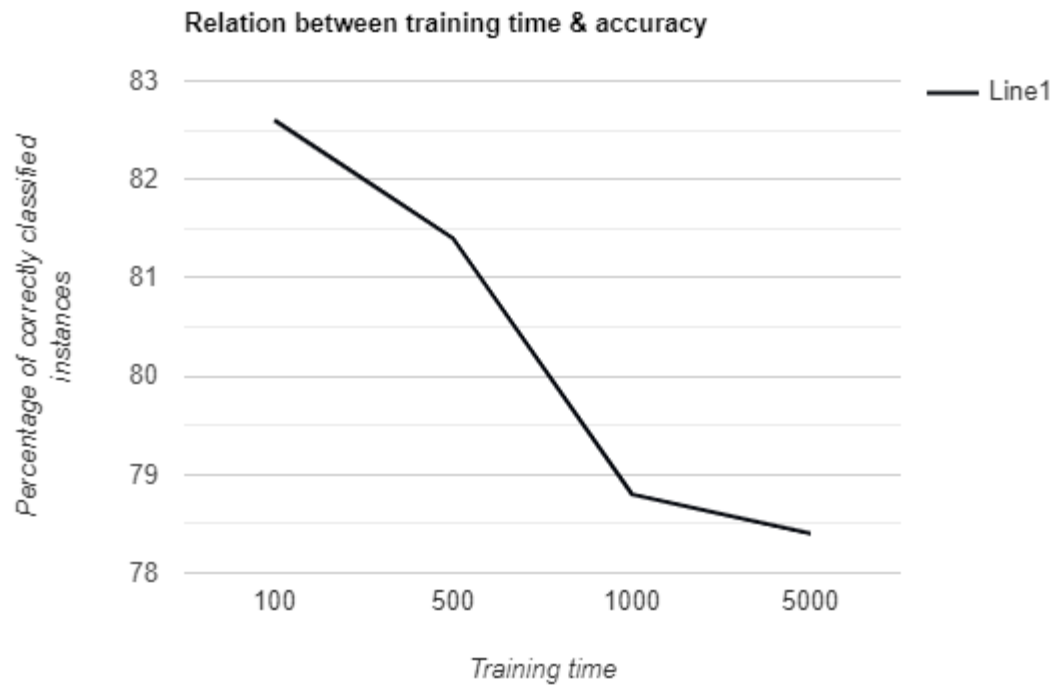
Precision: proportion of instances that are truly of a class divided by the total instances classified as that class

Recall: proportion of instances classified as a given class divided by the actual total in that class (equivalent to TP rate)

F-Measure: A combined measure for precision and recall calculated as $2 * \text{Precision} * \text{Recall} / (\text{Precision} + \text{Recall})$

6-

Training Time	Percentage of correctly classified instance
100	82.6398
500	81.9444
1000	78.8194
5000	78.4722



- 7- When we analyze the graph, we can easily see that the rate of correctly classified instances is decreasing while training time is increasing. The reason of this situation is overfitting.

Overfitting occurs when the analysis of classification tends to be too close or exact with given dataset. This means if training time is increased, the occurrence of overfitting is also increased.