Abstract: A set of classification models were tried Iris dataset to predict the Iris flower. Each model was fine-tuned on different settings using WEKA software to achieve the least incorrect classification percentage. For each best setting, the model was trained & tested using percentage split and 10 fold cross validation for comparison.

Classifiers		**Percentage Split (66% training)**		**10 Fold Cross Validation**	
		Correctly Classified (%)	Incorrectly Classified (%)	Correctly Classified (%)	Incorrectly Classified (%)
Naïve Bayes		94.1176 %	5.8824 %	96.6667 %	3.3333 %
Decision Tree	LMT	98.0392 %	**1.9608 %**	94 %	6%
	REPTree (seed =2)	92.1569 %	7.8431 %	94.6667 %	5.3333 %
	Random Forest	92.1569 %	7.8431 %	94 %	6%
	J48	96.0784 %	3.9216 %	96.6667 %	3.3333 %
k-NN		96.0784 %	3.9216 %	96.6667 %	3.3333 %
Perceptron		96.0784 %	3.9216 %	96.6667 %	3.3333 %
Multiple Layer Neural Network		98.0392 %	**1.9608 %**	98 %	2 %

Naïve Bayes: Three methods to tune the parameters were tried using NaiveBayes. The best model was achieved when 'kernel' estimator was used which gave the least 5 incorrect instances. **Decision Tree**: LMT, REPTree, Random Forest and J48 were fine tuned. The default LMT model gave the best model with 1.9% incorrect instances. **k-NN**: At k=12, ('k' was taken as square root of sample size - 150) the model gave the least (3.3%)incorrectly classified instances. **Perceptron**: Hidden layer 1 with single neuron, all sigmoid nodes, Learning Rate = 0.3, seed = 0 with normalized attributes match the best model criteria, with 5 incorrect instances at 10 fold cross validation. **Multiple Layer Neural Network**: Using with hidden layer of 3 neurons the model has 4 incorrect instances @ 2.667% and when hidden layer has 4 neurons and has incorrect instances @ 2%.

Classifiers	**Pros**	**Cons**
NaïveBayes	Several options of tuning (though WEKA class lacks Laplace / m-estimate options)	Dataset size matters. Larger the dataset slower the performance
Decision Tree	Several options of tuning options & parameters (multiple options in WEKA). Easy to interpret & gives control.	Tends to overfit
k-NN	Easy to implement once the k-value is known. Multiple linkage & distance methods could be tried	Requires deeper understanding of dataset to make the right choice of linkage & distance methods.
Perceptron	With sigmoid nodes the model converges as other models.	Hard to tune parameters.
Multiple Layer NN	Achieves high accuracy	Takes time even for small dataset; kind of a black box

Conclusion:

LMT decision tree and multiple neural network models with percentage split (@ 66%) gave the best models that has least incorrectly classified instances (1) at 1.96%.