## 2C - Comprehensive Comparison:

We do a comparative study and analysis of the following Machine Learning models-

- Fisher Linear Discriminant:
  The Fisher Linear Discriminant Analysis is a classification technique that identifies the linear combination of features that characterizes or separates two or more classes. The model fits a Gaussian density to each class, assuming that all classes share the same covariance matrix.

- Linear Perceptron:
  A Perceptron is a linear classification algorithm that makes its predictions based on a linear predictor function combining a set of weights with the feature vector. It uses the Gradient Descent Algorithm to update the weights based on misclassified points on each iteration.

- Naive Bayes:
  Naive Bayes algorithms are a set of supervised statistical classification machine learning algorithms based on the Bayes probability theorem. An important assumption made by Bayes theorem is that the value of a particular feature is independent from the value of any other feature for a given the class.

- Logistic Regression:
  Logistic regression is a machine learning algorithm for classification. In this algorithm, the probabilities describing the possible outcomes of a single trial are modelled using a logistic function.

- Artificial Neural Network:
  Artificial neural network is a machine learning technique used for classification problems. ANN is a set of connected input output network in which weight is associated with each connection. It consists of one input layer, one or more intermediate layer and one output layer. A Multi-Layer Perceptron is a supervised learning technique with a feed forward artificial neural network through back-propagation that can classify non-linearly separable data.

The models are imported from the sklearn libraries. The data is scaled using the pre-processing sklearn library, StandardScaler and a 7-fold cross validation is done for each model using the sklearn in built method cross_validate.
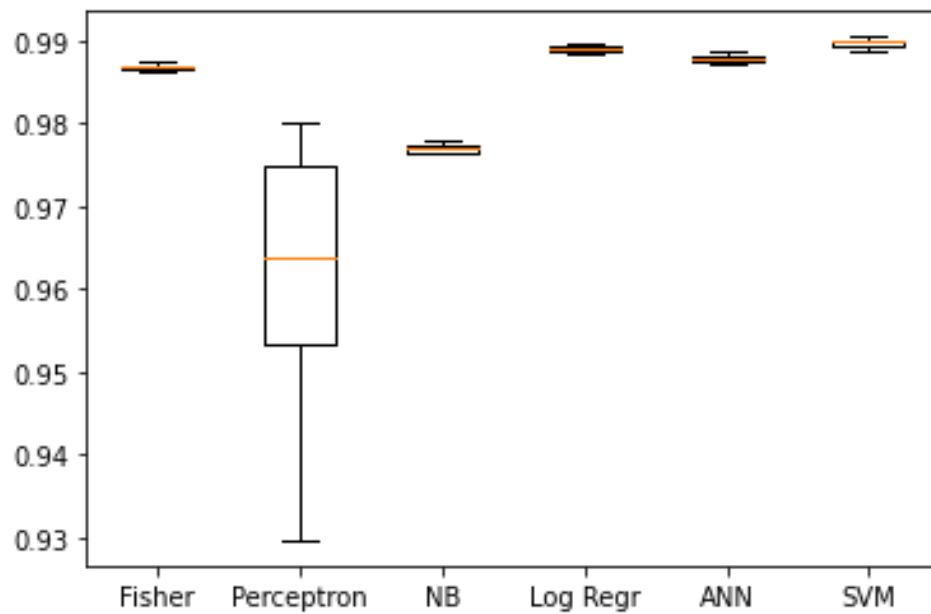
## Results:

| Model | Training Accuracy | Testing Accuracy |
|---|---|---|
| Fischer Linear Discriminant | 0.9867839788756373 | 0.9866922560063784 |
| Linear perceptron | 0.9614426214316038 | 0.960681417167005 |
| Naïve Bayes Classifier | 0.9769407035440894 | 0.9769040202001132 |
| Logistic Regression | 0.9890752402882411 | 0.9888919195100337 |
| Artificial Neural Networks | 0.9878562922294515 | 0.9875171933404873 |
| Support Vector Machine | 0.9897534582101397 | 0.9897717595034006 |

## Comparison :

Support Vector Machine has the highest and Linear Perceptron has the least testing Accuracy.

As the data here is inseparable so linear perceptron algorithm does not work as well, for inseparable data ANN and Support Vector Machine algorithms work better as it can be seen.

## Box Plots :

## Training Accuracies:

**Testing Accuracies:**