# BITS F464: MACHINE LEARNING

*Done by:*

| | |
|---|---|
| Rohan Maheshwari | 2017B4A70965H |
| Keshav Kabra | 2018AAPS0527H |
| Godhala Meganaa | 2017B3A70973H |


# 2C: **Comprehensive Comparison**

## Model Description

We designed a Classifier Class to instantiate each of the standard Classification Algorithms from sklearn. After that we performed 7-Fold Cross Validation to get Training and Testing Accuracies.

**1. Logistic Regression:**

- **Advantages:** Efficient, require low computational resources,easy to implement
- **Limitations:** Cannot solve non-linear Problems,probability of over fitting is high,unstable for well separated classes

**2.Naive Bayes:**

- **Advantages:** Performance is high if assumption of independence holds true,requires low train data
- **Limitations:** In real life, it's almost impossible to find predictors which are independent,zero probability for variable in test which hasn't occurred in train

**3.Linear Perceptron:**

- **Advantages:** Gives good results for binary classification on a linearly separable data
- **Limitations:** Data should be linearly separable

**4.Fisher Linear Discriminant:**

- **Advantages:** simple and effective method for classification
- **Limitations:** small sample size problem.

**5.Support Vector Machine:**

- **Advantages:** SVM is more effective in high dimensional spaces, and is memory efficient
- **Limitations:** Does not perform very well when the data set has more noise

**6.Artificial Neural Networks:**

- **Advantages:** Store information in the entire network, ability of parallel processing
- **Limitations:** ANNs can lead to unnecessarily long training times

Train Accuracies
( over 7-Fold Cross Validation )

Train Accuracies

| | LR | NB | LP | FLD | SVM | ANN |
|---|---|---|---|---|---|---|
| 0 | 0.986656 | 0.979085 | 0.983769 | 0.989286 | 0.991018 | 0.990762 |
| 1 | 0.986720 | 0.981716 | 0.983833 | 0.988773 | 0.990120 | 0.990569 |
| 2 | 0.986656 | 0.981844 | 0.983897 | 0.988773 | 0.990312 | 0.990312 |
| 3 | 0.987169 | 0.978700 | 0.983576 | 0.989350 | 0.990697 | 0.990377 |
| 4 | 0.986271 | 0.986784 | 0.983384 | 0.989607 | 0.990633 | 0.990569 |
| 5 | 0.987233 | 0.987875 | 0.984025 | 0.989735 | 0.990954 | 0.990762 |
| 6 | 0.986656 | 0.984347 | 0.983449 | 0.989158 | 0.990249 | 0.990826 |

Test Accuracies
( over 7-Fold Cross Validation )

Test Accuracies

| | FLD | LP | NB | LR | ANN | SVM |
|---|---|---|---|---|---|---|
| 0 | 0.987298 | 0.980754 | 0.983064 | 0.988838 | 0.988838 | 0.989992 |
| 1 | 0.986143 | 0.981909 | 0.982679 | 0.989992 | 0.990377 | 0.990762 |
| 2 | 0.988068 | 0.984988 | 0.982679 | 0.991147 | 0.992687 | 0.992302 |
| 3 | 0.983834 | 0.980370 | 0.984219 | 0.989222 | 0.989992 | 0.991532 |
| 4 | 0.989607 | 0.985758 | 0.984604 | 0.988068 | 0.989222 | 0.988838 |
| 5 | 0.984219 | 0.984988 | 0.982294 | 0.987683 | 0.989992 | 0.988838 |
| 6 | 0.987678 | 0.984213 | 0.985753 | 0.989218 | 0.989603 | 0.989988 |

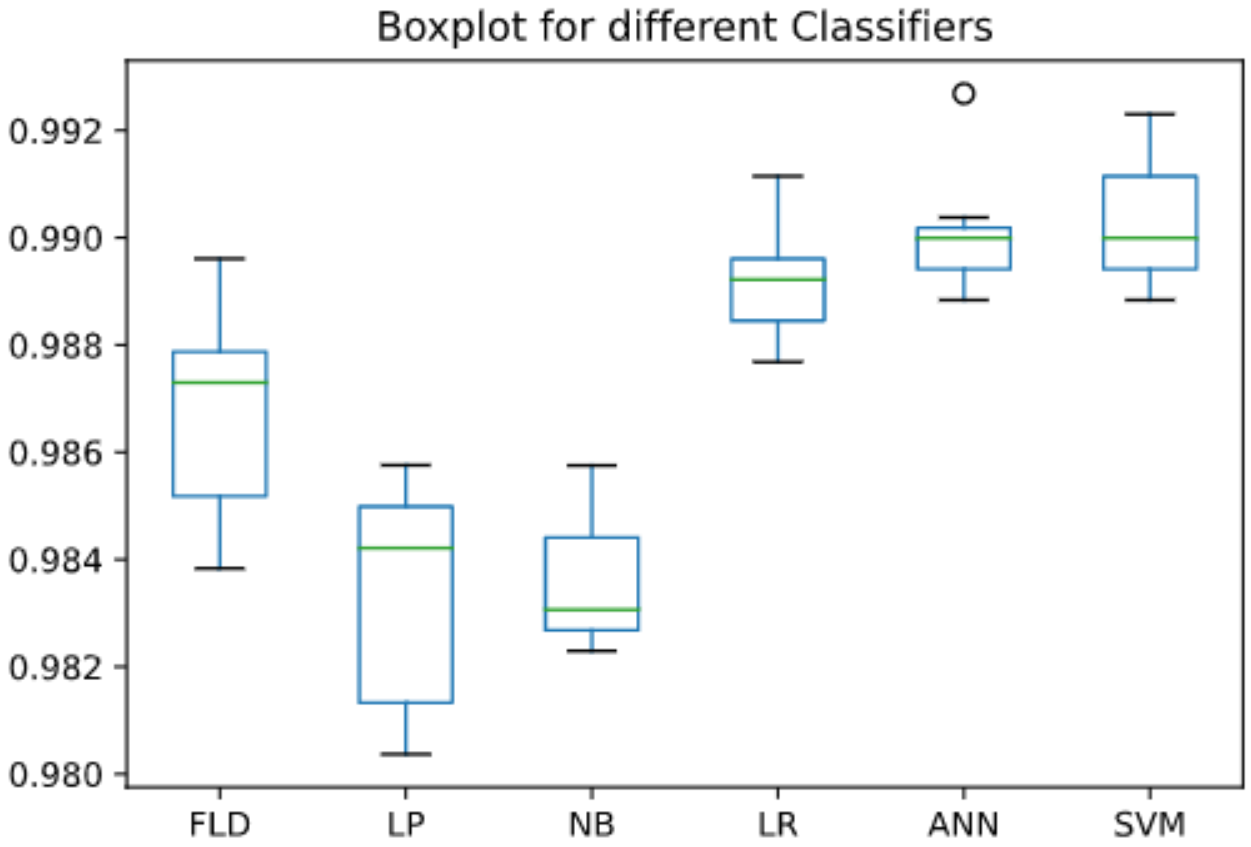
| Model | Avg Train Accuracy | Avg Test Accuracy | Avg MSE on Test |
|---|---|---|---|
| **FLD** | 0.986766 | 0.986692 | 0.013308 |
| **LP** | 0.982907 | 0.983283 | 0.016717 |
| **NB** | 0.983705 | 0.983613 | 0.016387 |
| **LR** | 0.989240 | 0.989167 | 0.010833 |
| **ANN** | 0.990569 | 0.990102 | 0.009898 |
| **SVM** | 0.990597 | 0.9903215 | 0.009678 |

The algorithm which performed **best** was **SVM** with an Average test accuracy of 96.11% and the algorithm which performed the **worst** was **LP**.

**Reasons:**

- From the results we observe that the SVM model outperformed the other models and reached the highest levels of accuracy on the validation sets. It is possible that after projecting the data into an infinite dimensional space using the radial basis function kernel, the transformed data contains an effective margin / decision boundary, using which the SVM can distinguish between both the classes easily.
- The model that performed the worst was perceptron since the model is derived on the assumption that the data is linearly separable which is not true in most of the cases so is a bad assumption to make. When the data given is not linearly separable we won't even know whether the perceptron is getting better on every iteration thereby after a large number of iterations leading to a reasonable performing classifier.

## Box Plot


Boxplot for different Classifiers
0.992
0.990
0.988
0.986
0.984
0.982
0.980
FLD
LP
NB
LR
ANN
SVM