**Machine Intelligence and Learning**

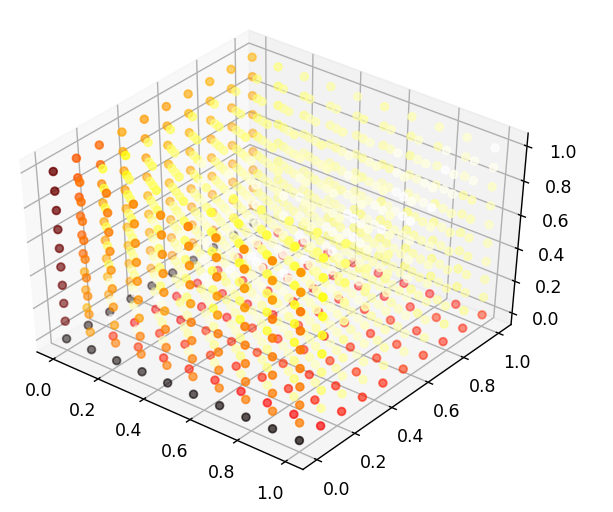
Indian Institute of Technology Delhi

COURSE ASSIGNMENT 2

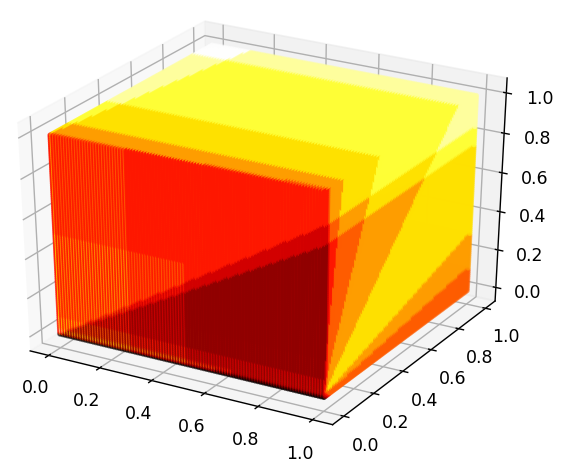
Mayank Mishra 2016EE30506 Chahat Chawla 2016MT10492

**Medical Dataset**

**Visualizations**

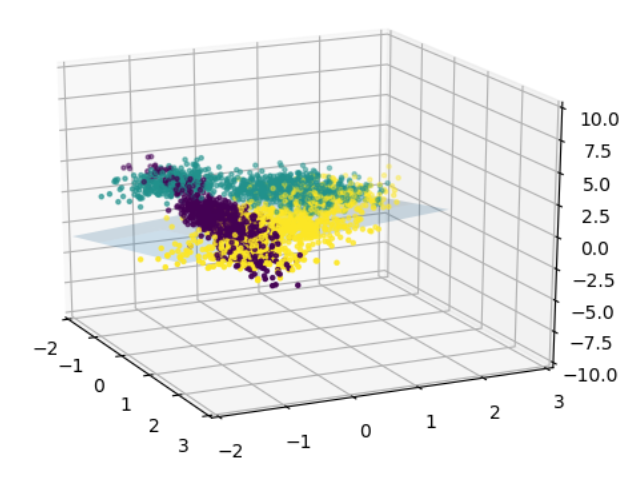


(a)

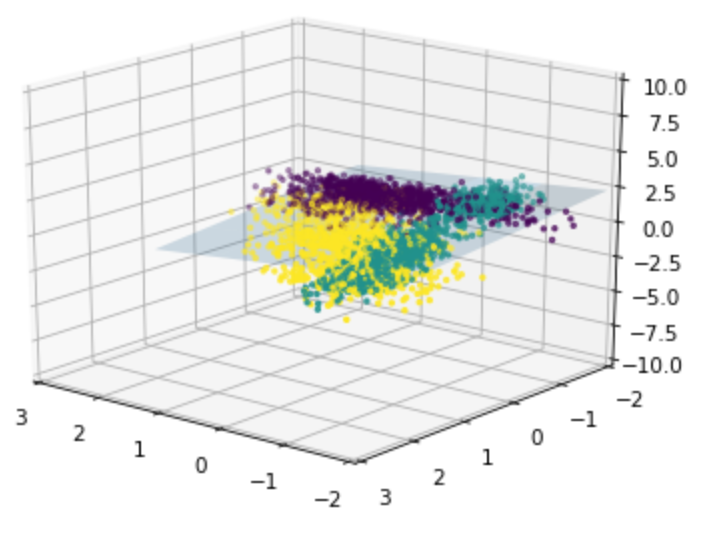


(b)

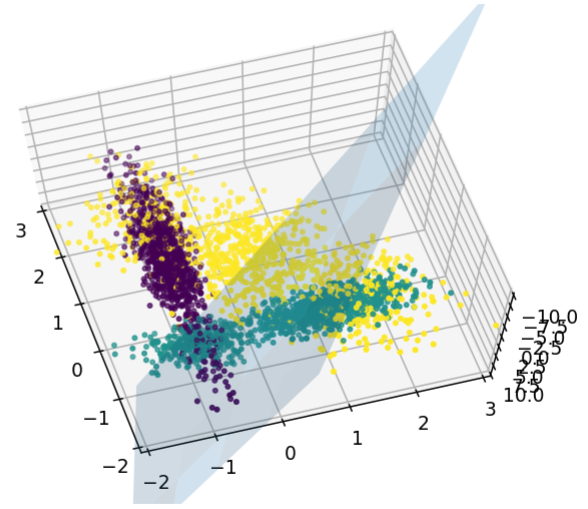
Fig. 1: Hyperparameter Tuning (Threshold vs Accuracy) [Color indicates the accuracy for the three classes]



**(a)**

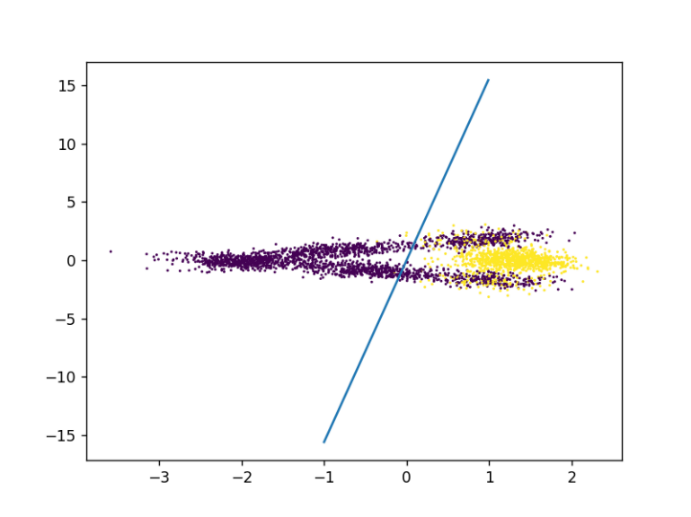


**(b)**

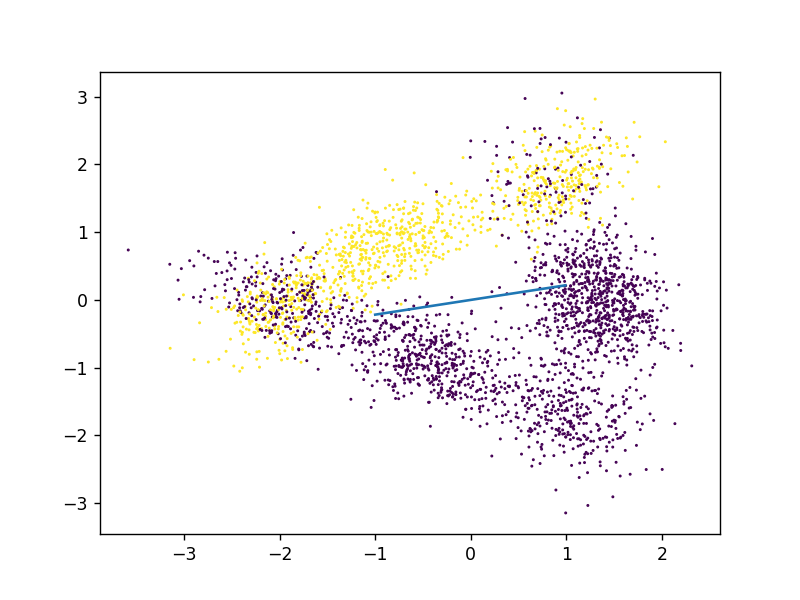


**(c)**

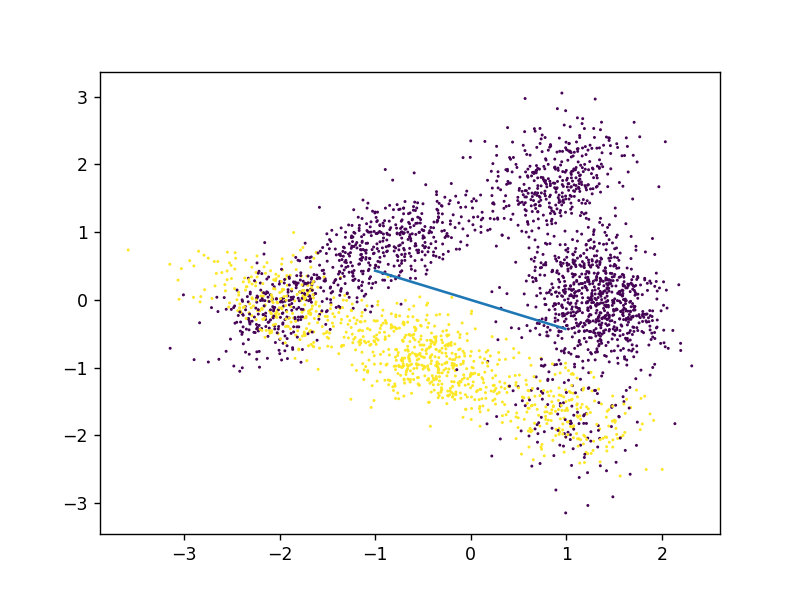
Fig. 2: Visualizing Hyperplanes learnt by Perceptron[(a) Medication (b) Healthy (c) Surgery]



**a)**

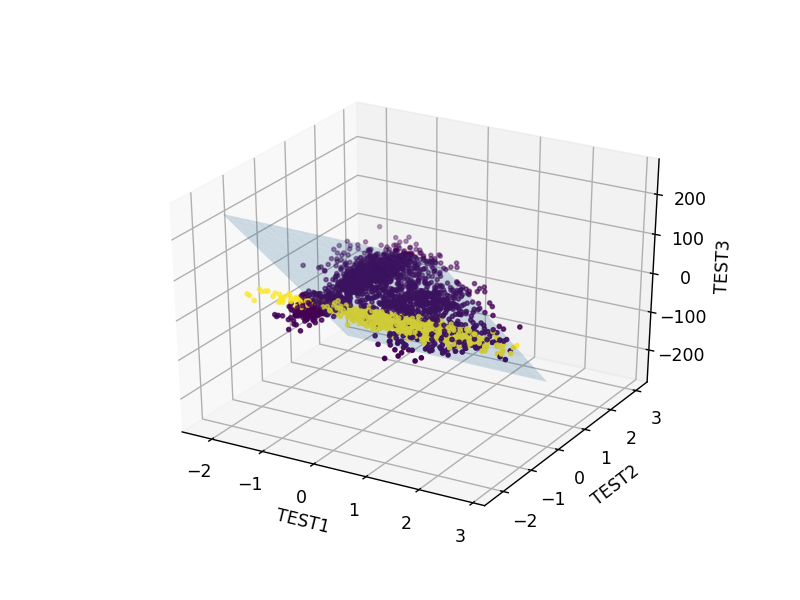


**b)**

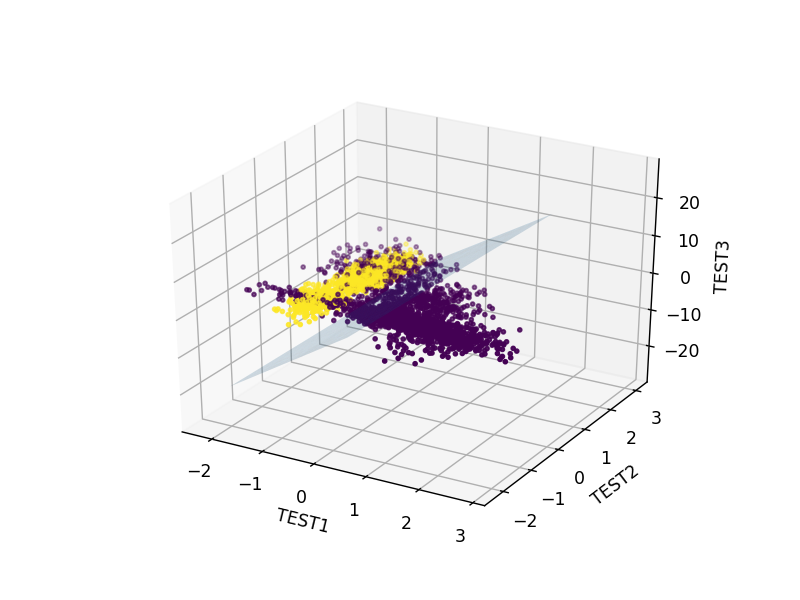


**c)**

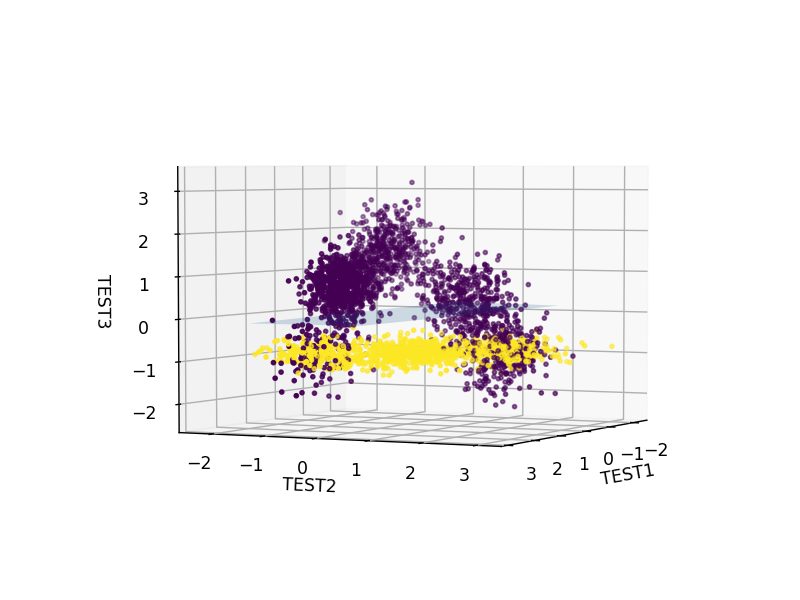
Fig. 3: Visualizing Hyperplanes learnt by Perceptron **with PCA** [(a) Medication (b) Healthy (c) Surgery]



a)



b)



**c)**

Fig. 4: One vs All Classification using FLDA [(a) Medication (b) Healthy (c) Surgery]

|  |  |
| --- | --- |
| **Algorithm** | **Accuracy** |
| Logistic Regression (Without Regularization) | 89.33% |
| Logistic Regression (With Regularization) | 88.93% |
| Softmax Classifier | 89.57% |
| Perceptron | 59.23% |
| Perceptron (after PCA) | 80.08% |
| SVM (Not Linearly Separable) |  |
| FLDA (multiclass: healthy, surgery, medication) | 77.56%, 81.67%, 80.77 % |

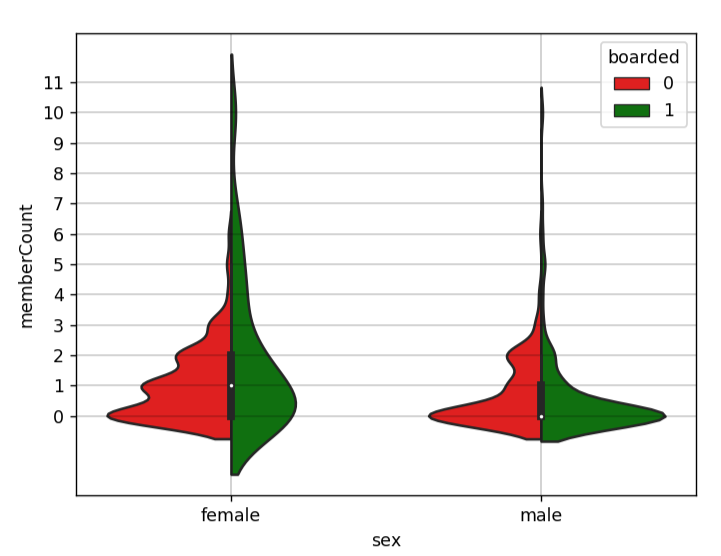
Table 1: Accuracy with different algorithms

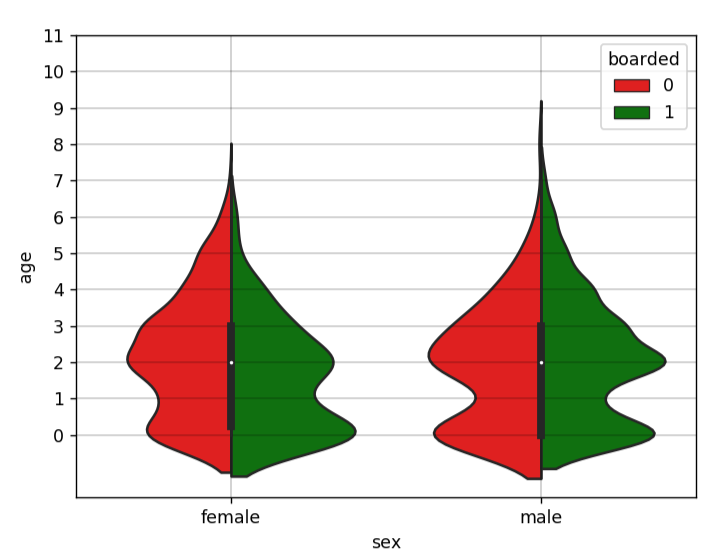
**Railway Dataset**

|  |  |
| --- | --- |
| **Algorithm** | **Accuracy** |
| Logistic Regression (Without Regularization) | 21.37% |
| Logistic Regression (With Regularization) | 21.37% |
| Softmax Classifier | 21.37% |
| Perceptron | 59.93% |
| SVM (Linearly Separable Assumption) | 59.93% |
| SVM (Not Linearly Separable) | 77.67% |
| FLDA | 50.00% |

Table 2: Accuracy with different algorithms

**Data Visualization**





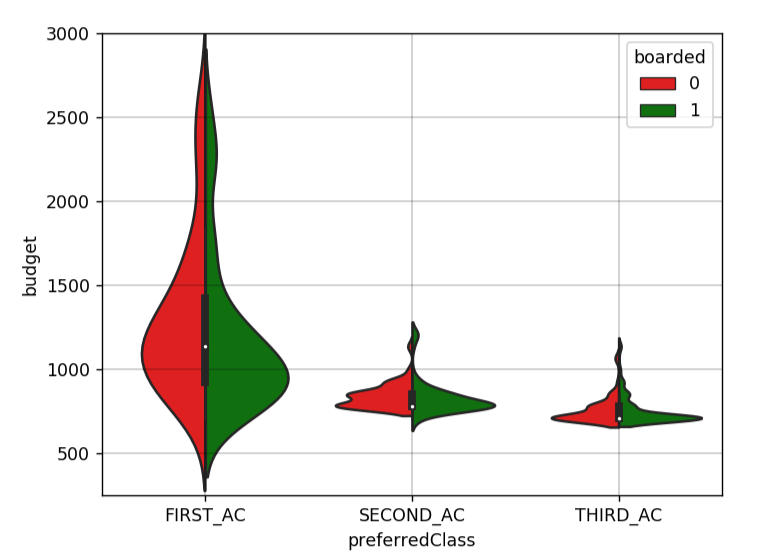


Fig. 5: Class Conditional Distribution of Features

**Fashion MNIST Dataset**

|  |  |
| --- | --- |
| **Algorithm** | **Accuracy** |
| Logistic Regression (Without Regularization) | 80.08% |
| Perceptron | 63.15% |

**River Dataset**

**Data Visualization**

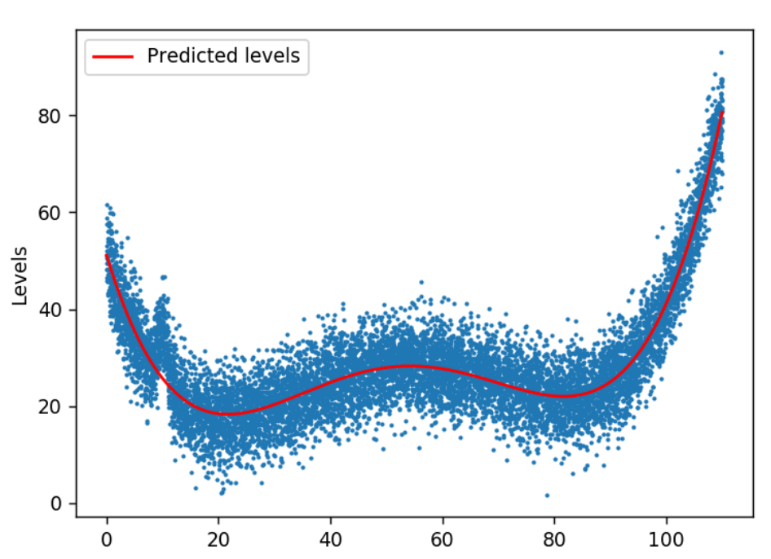


Fig.6: Original Data (in blue) and the predicted Curve (in Red)

Due to the nature of the problem, we apply Regression.

**Learnings from the Assignment**

* Gauging the performance of the algorithms designed for the linearly separable datasets on non-linearly separable datasets was one interesting aspect of the assignment.
* We apply transformation onto higher dimensional space to “improve linear separability” and thence aiding the learning of “better” separating hyperplanes.
* We learnt to make the optimal bias-variance tradeoff, by dividing the dataset into the Train/CV/Test Sets and tuning hyperparameters.
* Being a medical application, trading the different types of metrics (type-I, type-II errors) with medical dataset could be very much different. (An insight into the actual deployment of the Machine Learning models)
* We learnt to deal with Multiclass Classification tasks, via the One vs All and other approaches.
* Realized the significance of choosing the right Kernel for SVM. Specifically, we experimented with Polynomial, Gaussian, Sigmoidal, Laplace Kernels.
* We employed both the batch and stochastic versions of Perceptron and saw the difference while learning (due to Stochastic nature of the former) but to ultimate convergence.
* We employed 4-dimensional tensor to boost the performance of softmax classifier.

**Implementation Details**

* Local Optimums are attained with Softmax on Medical Dataset on different runs.
* Regularization did not help us in our case.
* We verified in working that Cross-Entropy Loss converges faster than Squared Error Loss.