Project Summary

Many machine learning algorithms are not interpretable. Hence the objective of this project is to create an interpretable algorithm which gives accuracy as good as the accuracy of other complex algorithms.

To achieve this, the algorithm mCLESS is developed using the least squares method.

Experiments are done on four data sets

1. Synthetic Data 1
2. Synthetic Data 2
3. Iris data
4. Wine data

Performances of following classifiers are then compared to the performance of mCLESS.

1. Logistic Regression
2. K-neighboursClassifiers
3. SVM - rbf
4. RandomForestClassifier

The performance meteric used for the experiments is percentage accuracy.

All Algorithms are run 100 times, creating random 70:30 split of the data for training and testing in each run. Then the average accuracy is reported for each algorithm. From these experiments , we found that that mCLESS gives accuracy as good as other classifiers at a much less computational cost.

Feature expansion -

Feature expansion is done by adding a dimension equal to distance of a data sample from a particular point p in the dataset.

The startegies implemented to choose p are as follows:

1. Visually choose the point p for synthetic data1 and synthetic data2.
2. Chose the center of the dataset as point p.
3. Linesearch for point p along the least sqaures quadratic polynomial.
4. Linesearch along the triangle formed by the centers of the cluster.

The center of the dataset can be chosen as p because when we add the dimension, all points would go away from center, thus increasing the spaces between the clusters. It can make classses more separable.

For Synthetic data 1, the exepriments were done by chosing p as the center of the data, by linesearch, and by linesearch of the triangle. It was found that accuracy for this dataset remains almost the same, irrespective of the choice of point p.

For Synthetic data 2, the centers of the class clusters are almost colinear. If the point p is chosen as the center of the class 1 cluster then the distances of the points in other class clusters from p are comparatively larger than the distances of points in class 1 cluster. Thus when we add $||\mathbf{x} - \mathbf{p}||$ as the additional dimension with p chosen to be the class 1 center, the points in other classes move away with respect to the class 1 cluster in the higher dimensional space. This breaks the colinearity of the clusters and makes them more separable.

The experiments validate this point. In all the four strategies mentioned above, the accuracy increased from around 71% to around 95% when the point p is the center of the cluster of class 1 in the Synthetic dataset 2. Coincidently, the overall center of the entire data set is also near the center of the class 1 cluster. Hence, its accuracy also showed similar improvement.

For Iris data, by using p as the overall center of the entire data set, the accuracy improved from 82.96% to 95.41%. For doing linesearch for Iris data, we chose the feature at the index 2 and 3 because the classes looked more separable in the pairplot of features at 2 and 3. The accuracy can be futher improved to 97.43% by the point $p = [1.0, 0.04]$ which was obtained using the linesearch.

For Wine dataset, the accuracy before adding dimension is 98.61%. The accuracy of wine data did not improve much by trying the above mentioned strategies to find p.

Conclusion -

I have successfully implemented and tested the algorithm mCLESS for multi-class classification, which is also interpretable. The parametres of the classifier are visualized and it showcases why the classifier is interpretable. The weight matrix calculated by the least squares method represents the family of lines for each class such that if the distance of a point from the line $L_j = 0$ is maximum, then the point belongs to class j.

For the datasets which cannot be separated using one versus all techniques, it appears that the drop is accuracy is because the decision boundaries (hyperplanes) become almost parallel and collinear to one another. This project shows that a significant improvement in the prediction accuracy can be made by adding an additional dimension to the data by the equation: $||\mathbf{x} - \mathbf{p}||$, where the point $\mathbf{p}$ is chosen such that it breaks collinearity and facilitates clear separation of the classes. Making the data more separable increases the accuracy of any classification algorithm. It was found that $\mathbf{p}$ can be chosen to be in the region near the center of the overall data set. It improves the accuracy for colinear clusters, and also does not significantly affect the accuracy in cases where the classes are not strongly non-colinear.