Assignment 3: Linear Classifiers(ML)

Name:Satya Prakash Sharma
B17CS048

Dr. Gaurav Harit

For Linear classification, I have chosen the dataset of heart diseases patients for predicting the possibilities of heart attacks. The dataset has 14 numeric attributes.

For the linear classification, I have performed different algorithms for predicting the possibilities of heart attack.
The algorithm that I have choose are as follows:

1. Half Space(using sklearn perceptron function)
2. Logistic Regression(using sklearn inbuilt function)
3. SVM classifier(using a linear kernel)
4. SVM classifier(using a Polynomial kernel)
5. SVM classifier(using a Gaussian(radial basis) kernel)
6. Logistic regression(using SGD procedure)

The Accuracy for the above algorithm are as follows:

Classifier Name	**Accuracy(70:30)**	**Accuracy(80:20)**	**Accuracy(90:10)**
Half Space	73.63%	49.18%	70.97%
Logistic Regression	81.32%	88.52%	83.87%
SVM classifier(using linear kernel)	81.32%	86.89%	83.87%
SVM classifier(using a polynomial kernel)	83.52%	90.16%	83.87%
SVM classifier(using a Gaussian(radial basis) kernel)	82.42%	86.89%	80.65%
Logistic regression(using SGD procedure)	84.62%	86.89%	83.87%

From the above experiment, I conclude that:

Half Space Classifier/Perceptron

- Half Space Classifier gives less accuracy from all other algorithms. Logistic regression performs better than Half space classifier. Since Logistic regression uses logistic function and the half space classifier uses a perceptron algorithm which internally uses step function.
- The parameter vector for the perceptron may be arbitrarily scaled compared to the one derived by logistic regression. Any scaling of the parameter vector will define the same boundary, but the probabilities calculated by logistic regression depend on the exact scaling.
- The output from a step function can of course not be interpreted as any kind of probability.
- Since a step function is not differentiable, it is not possible to train a perceptron using the same algorithms that are used for logistic regression.
- This algorithm is good for the data splits, 70:30 and 90:10. In 70:30 it is somewhat less and seems to be overfitting.

Logistic regression

- Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. In regression analysis, logistic regression (or logit regression) is estimating the parameters of a logistic model (a form of binary regression). The hypothesis in logistic regression provides a measure of uncertainty in the occurrence of a binary outcome based on a linear model. The output is bounded asymptotically between 0 and 1, and depends on a linear model.
- This algorithm is good for all the data splits, 70:30, 80:20 and 90:10. In 80:20 it is somewhat less and seems to be overfitting.

SVM Classifier(Linear kernel)

- SVM can be line/plane or d-1 space for d dimensional data which helps us to segregate the binary data.
- It segregates the binary data in 2 sets: first conclude yes and other for no.
- The regularization parameters are used to control the weightage we want to give to the w vector and loss(errors) accordingly.
- The linear kernel(usually faster than other kernels) is used to simplify and approx the dot products for computational advantage.
- This algorithm is good for all the data splits, 70:30, 80:20 and 90:10. In 80:20 it is somewhat less and seems to be overfitting.

SVM Classifier(Polynomial kernel)

- Same as SVM, here we choose a polynomial kernel.
- a polynomial kernel is better in approximation as compared to the linear kernel.
- Less computational than the gaussian kernel, but more than the linear one.
- It can be seen from the Accuracy.
- This algorithm is good for all the data splits, 70:30, 80:20 and 90:10. In 80:20 it is somewhat less and seems to be overfitting.

SVM Classifier(Guassian kernel)

- Same SVM, here we choose a gaussian kernel(radial basis kernel).
- The Gaussian kernel gives more accuracy than linear and polynomial. Which can be seen from the accuracy above.
- It is more computationally expensive linear and polynomial.
- This algorithm is good for all the data splits, 70:30, 80:20 and 90:10. In 80:20 it is somewhat less and seems to be overfitting.

Logistic regression using SGD

- Stochastic Gradient Descent is an efficient approach for fitting classifiers under convex loss functions such as Logistic Regression and SVMs,
- We can also use them for non differentiable functions with their partial differentiation vectors.
- This algorithm is good for all the data splits, 70:30, 80:20 and 90:10. In 80:20 it is somewhat less and seems to be overfitting.