CS 5011: Assignment 2 report

Due on Thursday, October 16, 2014

Sure Manoj Kumar

CS12B028

Question 1.

The parameters are chosen based on the accuracy scores.
Linear : 61.43% accuracy

polynomial:
maximum accuracy - 68.58% for degree=2 and gamma=2.

guassian(rbf):
maximum accuracy - 69.68% for gamma = 2 and C = 4 .

sigmoid:
maximum accuracy - 54.37% for gamma=1 and coef0=0 .

Question 2.

For the original back-propagation algorithm,the performance measures are:
parameters used $\gamma_r = 0.01$ and iterations=200.
Accuracy = 0.6125
Precision for coast = 1.0
Precision for forest = 0.90
Precision for insidecity = 0.54
Precision for mountain = 0.58
Recall for coast = 0.05
Recall for forest = 0.5
Recall for insidecity= 1.0
Recall for mountain = 0.9
F_measure for coast = 0.095
F_measure for forest = 0.64
F_measure for insidecity = 0.7017
F_measure for mountain = 0.7058
Updated rules for the new error function :

$$\beta_{km}^{(r+1)} = \beta_{km}^{(r)} - \gamma_r(\sum_{i=1}^{N}(\frac{\partial R}{\partial \beta_{km}^{(r)}}) + 2\gamma\beta_{km}^{(r)})$$

$$\alpha_{ml}^{(r+1)} = \alpha_{ml}^{(r)} - \gamma_r(\sum_{i=1}^{N}(\frac{\partial R}{\partial \alpha_{ml}^{(r)}}) + 2\gamma\alpha_{ml}^{(r)})$$

parameters used $\gamma_r = 0.01$ and iterations=100 and $\gamma = 0.01$.
Accuracy = 0.6875
Precision for coast = 0.6
Precision for forest = 0.73
Precision for insidecity = 0.75
Precision for mountain = 0.625
Recall for coast = 0.3
Recall for forest = 0.95
Recall for insidecity= 0.75
Recall for mountain = 0.75
F_measure for coast = 0.4
F_measure for forest = 0.82
F_measure for insidecity = 0.75
F_measure for mountain = 0.68

parameters used $\gamma_r = 0.01$ and iterations=100 and $\gamma = 0.1$.
Accuracy = 0.6625
Precision for coast = 0.63
Precision for forest = 0.65
Precision for insidecity = 0.66
Precision for mountain = 0.68
Recall for coast = 0.35
Recall for forest = 0.95
Recall for insidecity= 0.6
Recall for mountain = 0.75
F_measure for coast = 0.45
F_measure for forest = 0.77
F_measure for insidecity = 0.63
F_measure for mountain = 0.71

parameters used $\gamma_r = 0.01$ and iterations=100 and $\gamma = 1$.
Accuracy = 0.5625
Precision for coast = 0.55
Precision for forest = 0.62
Precision for insidecity = 0.47
Precision for mountain = 0.57
Recall for coast = 0.25
Recall for forest = 0.9
Recall for insidecity= 0.55
Recall for mountain = 0.55
F_measure for coast = 0.34
F_measure for forest = 0.73
F_measure for insidecity = 0.51
F_measure for mountain = 0.56

parameters used $\gamma_r = 0.01$ and iterations=50 and $\gamma = 10$.
Accuracy = 0.5375

Precision for coast = 0.41
Precision for forest = 0.62
Precision for insidecity = 0.48
Precision for mountain = 0.57
Recall for coast = 0.25
Recall for forest = 0.75
Recall for insidecity= 0.6
Recall for mountain = 0.55
F_measure for coast = 0.31
F_measure for forest = 0.68
F_measure for insidecity = 0.53
F_measure for mountain = 0.56

parameters used $\gamma_r = 0.001$ and iterations=25 and $\gamma = 100$.
Accuracy = 0.3
Precision for coast = 0.21
Precision for forest = 0
Precision for insidecity = 0.31
Precision for mountain = 0.33
Recall for coast = 0.25
Recall for forest = 0
Recall for insidecity= 0.20
Recall for mountain = 0.30
F_measure for coast =0.22
F_measure for forest = -
F_measure for insidecity = 0.30
F_measure for mountain =0.31

The new error function is similar to adding ridge in ridge regression.The advantage of this error function is numerical values of weights will be low as compared to previous weights which were derived from original back propagation algorithm.This is because,we have added weights also in the error term ,which makes the numerical values of weights to be low.The more the value of γ the less the values of weights.and as γ increases the accuracy/performance decreases.As the value of the gamma increases,the function will try to minimise the weights rather than error caused due to mis-classification which makes the accuracy to drop as the value of gamma increases.

Question 3.

For L2 logistic regression, the performance measures are
Accuracy - 0.95
Precision for forest class - 1.0
Precision for mountain class - 0.909090909091
Recall for forest class - 0.9
Recall for mountain class - 1.0
F-measure for forest class - 0.947368421053
F-measure for mountain class - 0.952380952381

For L1 logistic regression using Boyd's code ,the performance measures are $for\ \lambda = 0.01$

Accuracy - 0.975
Precision for forest class - 1.0
Precision for mountain class - 0.952380952381
Recall for forest class - 0.95
Recall for mountain class - 1.0
F-measure for forest class - 0.974358974359
F-measure for mountain class - 0.975609756098

Question 4.

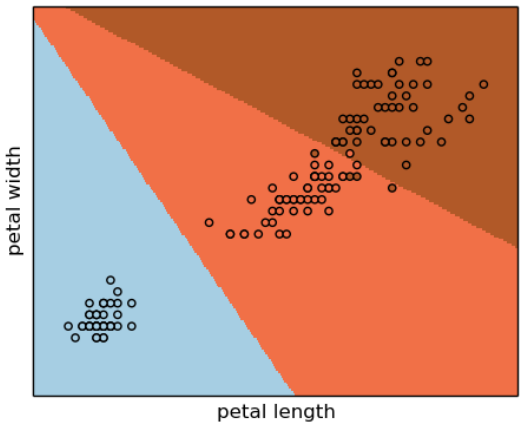

Figure 1: LDA

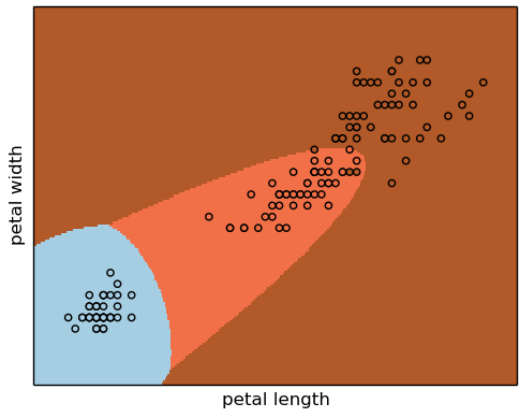

Figure 2: QDA

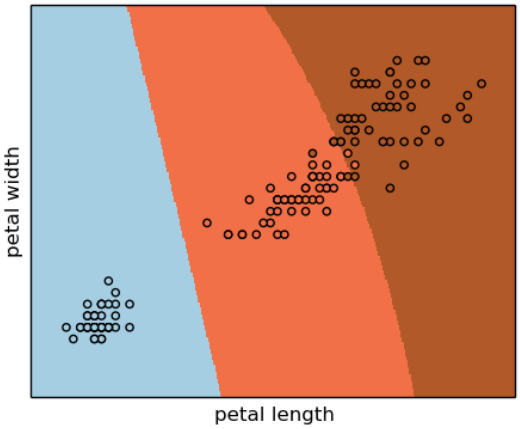

Figure 3: RDA with regparam as 0.3