Programming Assignment-02
Intro to Machine Learning -CS5011

Sunil Kumar J S(NA13B031)

October 15, 2016

1 Support Vector Machines

1.1 Linear-kernel

Best parameters: C-value= 0.1

1.2 Polynomial-kernel

Best parameters: degree=2 and C=0.8

1.3 Gaussian-kernel

Best parameters: γ=0.01, C=1

1.4 Sigmoid-kernel

Best parameters: C=0.9 , γ=0.005 , coefficeint=-0.5

2 Neural Network

2.1 Performance of NN without regularizing weights

Label	Precision	Recall	F1-score
mountain	0.83	0.25	0.38
forest	0.55	0.75	0.63
insidecity	0.54	0.85	0.55
coast	0.68	0.55	0.61

2.2 Update Rule with regularization of weights

Following changes have to be made when weights are regularized. Notations are same as used in class.These changed derivatives have to be plugged in backpropagation equation.

$$\frac{\partial R_i}{\partial \beta_{km}} = \delta_{ki} z_{mi} + 2\gamma\beta_{km} \tag{1}$$

$$\frac{\partial R_i}{\partial \alpha_{ml}} = s_{mi} x_{il} + 2\gamma\alpha_{ml} \tag{2}$$

2.3 Effect γ in results

γ acts analogous to λ in Ridge Regression.Larger values of γ shrink weights. towards zero.This improves the prediction as only those networks whose weights are more are considered.In a way this selects best features. As expected larger values of γ makes α and β sparser.**For given Dataset γ=1 and 10 yielded best results**

2.4 Performace results with regulurization

Best results were obtained when γ is 10.

2.4.1 Performance when γ=10

Label	Precision	Recall	F1-score
mountain	0.83	0.25	0.38
forest	0.55	0.75	0.63
insidecity	0.54	0.85	0.55
coast	0.68	0.55	0.61

2.4.2

For other values of γ , following accuracies and F1- scores were obtained

γ	F1-score	Accuracy
.01	0.34	0.31
1	0.59	0.59
10	0.64	0.60
100	0.49	0.41

3 Decision-Trees

3.1 Performace report

Following results are obtained upon running j48 Decision tree with default parameters (minNumobj=2)

Metric	Decision-Tree		
	class p	class e	overall
Precision	1.0	1.0	1.0
Recall	1.0	1.0	1.0
f-measure	1.0	1.0	1.0

Table 1: parameters considered:default parameters according to weka

3.2 Effect of minNumobj on performace

As minNumobj is increased the **performance** of the classifier goes down a little, however **interpretability** of tree increases i.e no of leaves and size of tree decreases significantly. For example when minNumobj is set to **1500**, the size of tree is just **10** against 30 (minNumobj=0) and very little performance is compromised. (F1-score is **0.994** when minNumobj=1500).So **minNumobj allow us to tradeoff between Interpretability and performance**

3.3 What does minNumobj mean?

minNumobj sets minimum instances per leaf.This guarantees that at each split, atleast two of the branches will have the instances atleast minNumobj.Separating one instance from 1000 instances doesn't give us much information, so minNumobj sets a minimum amount of separation. However, if

minNumobj	size of Tree	Weighted F1-score
2	30	1
25	24	0.994
1500	10	0.994
3000	3	0.488

Table 2: Tradeoff between interpretability and performace

we have a node with four branches, and two of them end up with 0 instances, the other two with 50 each, the branching still produced information.

3.4 what happens on doing reducederror Pruning

Reduced pruning error reduced the number of leaves and size of tree increasing accuracy or keeping it same. For example:
Case-1:minNumobj=30, without reduced pruning, **no. of leaves =21,size of Tree=25**,f1-score=.994
Case-2:minNumobj=30,with reduced pruning , **no of leaves =17,size of tree =19**,f1-score=0.994

3.5 Important features in deciding whether a mushroom is edible or not

- odor
- spore-print color
- gill-size
- cap-surface
- gill -spacing
- population