CSE 574

Machine Learning

Prof: Dr. S Srihari

Project 1

Jruvika Bhimani

UBitName: jruvikam

UBitNumber: 50207613

(1) Data Partition:

The LeToR data has been partitioned using the slicing function in python to include:
- 80% of the records in Training set
- 10% in validation
- Remaining 10% in Test set.

The same has been followed in Slicing of Synthetic set.

Set	Start	End
Training Set	0	floor(0.8 * total#of_records)
Validation Set	floor(0.8 * total#of_records)+1	floor(0.8 * total#of_records) + floor(0.1* total#of_records
Testing Set	floor(0.8 * total#of_records) + floor(0.1* total#of_records) +1	End of File

(2) Hyper-parameter Tuning:

Hyper-Parameter	Value	Reason
M	10	Higher M the computational cost of the program increases and renders the processing infeasible.
λ	1	Used to avoid overfitting of the regression model.
μj	K-means Clustering	Using the K-means clustering method. μ is calculated for every cluster where number of clusters = M, and μ is considered as a cluster centre.
Σj	Identity Matrix	Best optimal value received with the variance as an Identity matrix.
$\eta(\tau)$	0.01	Best optimal value received.

(3) **Evaluation and Results:**

Linear Regression:

After getting μ for M clusters the μj and xn vectors are considered and to calculate the the feature vector i.e. ϕj:

$$\phi_j(\mathbf{x}) = \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_j)^\top \Sigma_j^{-1}(\mathbf{x}-\boldsymbol{\mu}_j)\right)$$

each ϕj then helps in creating the ϕ matrix:

$$\mathbf{\Phi} = \begin{bmatrix} \phi_0(\mathbf{x}_1) & \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \cdots & \phi_{M-1}(\mathbf{x}_1) \\ \phi_0(\mathbf{x}_2) & \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \cdots & \phi_{M-1}(\mathbf{x}_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_0(\mathbf{x}_N) & \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & \cdots & \phi_{M-1}(\mathbf{x}_N) \end{bmatrix}$$

this ϕ matrix is used to calculate the Closed-Form solution using:

$$\mathbf{w}^* = (\lambda\mathbf{I} + \mathbf{\Phi}^\top\mathbf{\Phi})^{-1}\mathbf{\Phi}^\top\mathbf{t}$$

where λ = 1 i.e.:

$$\mathbf{w}_{\mathrm{ML}} = (\mathbf{\Phi}^\top\mathbf{\Phi})^{-1}\mathbf{\Phi}^\top\mathbf{t}$$

wML is the CLOSED FORM solution for the Linear Regression.

$$y(\mathbf{x}, \mathbf{w}) = \mathbf{w}^{\top} \boldsymbol{\phi}(\mathbf{x})$$

Where y(x,w) is the target value. Hence using the below formula for E-RMS, we get the Root Mean Square Error for the Linear Regression:

$$E_D(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^{N} \{t_n - \mathbf{w}^{\top} \boldsymbol{\phi}(\mathbf{x}_n)\}^2$$

Substituting E(w*) in the below formula we get E-RMS:

$$E_{\mathrm{RMS}} = \sqrt{2E(\mathbf{w}^{\star})/N_{\mathrm{V}}}$$

Stochastic Gradient Descent:

We use the same Hyper-parameters set in Linear Regression model:

We calculate the E-RMS using:

$$\nabla E_D = -(t_n - \mathbf{w}^{(\tau)\top}\boldsymbol{\phi}(\mathbf{x}_n))\boldsymbol{\phi}(\mathbf{x}_n)$$

where tn is the target vector, w(τ) is the weight vector per iteration. We finalize the weight vector when the value of E-RMS converges i.e. become steady over iterations:

$$\nabla E = \nabla E_D + \lambda \nabla E_W$$

$$\Delta \mathbf{w}^{(\tau)} = -\eta^{(\tau)} \nabla E$$

$$\mathbf{w}^{(\tau+1)} = \mathbf{w}^{(\tau)} + \Delta \mathbf{w}^{(\tau)}$$

E-RMS Values:

LeToR Data

LeToR Data	Linear Regression	Stochastic Gradient Descent
Training Set	0.560987290594	0.561797338851
Validation Set	0.552230095624	0.569564863836
Testing Set	0.636461354255	0.648100515066

Synthetic Data

Synthetic Data	Linear Regression	Stochastic Gradient Descent
Training Set	0.631698174236	0.634496455696
Validation Set	0.647220572422	0.783426560713
Testing Set	0.631667785422	0.787848505967

Results:

UBitName = jruvikam
personNumber = 50207613

################## LETOR ##################

***************** Training Set *****************
M: 10.000000
lambda: 1.000000
eta: 0.010000
sigma: Identity matrix
wML:
[-0.10038044 0.41657243 0.22322325 0.12176405 0.73032927 0.37536641
0.41103069 -0.75008745 0.76397962 0.53542]
ERMS:
0.560987290594
w SGD:
[-0.10064633 0.3694492 0.29646184 0.09353877 0.753612 0.29289928
0.31587552 -0.71097692 0.76093895 0.48741239]
Erms SGD:
0.561797338851

***************** Validation Set *****************
M: 10.000000
lambda: 1.000000
eta: 0.010000
sigma: Identity matrix
wML:
[-0.10038044 0.41657243 0.22322325 0.12176405 0.73032927 0.37536641
0.41103069 -0.75008745 0.76397962 0.53542]
ERMS:
0.552230095624
w SGD:
[-0.04148487 0.48956662 0.26020592 0.18144553 0.33576551 0.59553955
0.23399934 0.11752796 0.46204152 0.3168952]

Erms SGD:
0.569564863836

****************** Testing Set ******************
M: 10.000000
lambda: 1.000000
eta: 0.010000
sigma: Identity matrix
wML:
[-0.10038044 0.41657243 0.22322325 0.12176405 0.73032927 0.37536641
 0.41103069 -0.75008745 0.76397962 0.53542   ]
ERMS:
0.636461354255
w SGD:
[-0.17110473 0.34591523 0.35547722 0.5473034   0.23266415 0.41938586
 0.21036499 0.02906629 0.6965924   0.37328071]
Erms SGD:
0.648100515066

################## SYNTHETIC DATA ##################

****************** Training Set ******************
M: 10.000000
lambda: 1.000000
eta: 0.010000
sigma: Identity matrix
wML:
[0.78637378 3.10847465 -2.32084459 -0.70811311 -1.07282424 -0.56594611
 1.16715306 -0.39506556 -0.42623686 2.07396753]
ERMS:
0.631698174236
w SGD:
[0.82464205 3.13227422 -2.31820635 -0.66442851 -1.05351146 -0.59292191
 1.17225794 -0.39842426 -0.42923789 2.07361962]
Erms SGD:
0.634496455696

```
***************** Validation Set *****************
M: 10.000000
lambda: 1.000000
eta: 0.010000
sigma: Identity matrix
wML:
[ 0.78637378  3.10847465 -2.32084459 -0.70811311 -1.07282424 -0.56594611
  1.16715306 -0.39506556 -0.42623686  2.07396753]
ERMS:
0.647220572422
w SGD:
[ 0.21268556  0.35188175  0.07111355 -0.03476025 -0.37067944  0.45854719
  0.06480144  0.19776766  0.11551296  0.20524059]
Erms SGD:
0.783426560713

***************** Testing Set *****************
M: 10.000000
lambda: 1.000000
eta: 0.010000
sigma: Identity matrix
wML:
[ 0.78637378  3.10847465 -2.32084459 -0.70811311 -1.07282424 -0.56594611
  1.16715306 -0.39506556 -0.42623686  2.07396753]
ERMS:
0.631667785422
w SGD:
[ 0.05306917 -0.09282635  0.36454512  0.1716028   0.29113859  0.04958528
  0.01854001 -0.04095363 -0.19328851  0.65919834]
Erms SGD:
0.787848505967
EXIT
```