

Cover page for answers.pdf
CSE353 Fall 2020 - Machine Learning - Homework 4

Your Name: Cynthia Cheung

Solar ID: 11756494

NetID email address: cynthia.cheung@stonybrook.edu

Names of people whom you discussed the homework with:

Aggelina Chatziagapi

Ajay Gopal Krishna

Peter Hwang

1.1

a. Linear SVM

Train Set:

Accuracy: 0.8146248579589079

Confusion matrix: [[24078 642]
[5394 2447]]

Test Set:

Accuracy: 0.8134635464652048

Confusion matrix: [[12088 347]
[2690 1156]]

b. RBF Kernel

Train Set:

Accuracy: 0.855962654709622

Confusion matrix: [[23407 1313]
[3377 4464]]

Test Set:

Accuracy: 0.8511762176770469

Confusion matrix: [[11757 678]
[1745 2101]]

The RBF kernel SVM classifier has a higher accuracy score than the linear SVM classifier for both the train and test set. This could mean that for this data set, the RBF kernel is a better SVM classifier. The train set for the RBF kernel SVM classifier has a higher accuracy score than the test set for the RBF kernel SVM classifier. Similarly, the train set for the linear SVM classifier has a higher accuracy score than the test set for the linear SVM classifier. This could be a result of an overfitting of data points as the classifiers are trained on the train set. However, the accuracy score of the RBF kernel SVM test set and the accuracy score of the RBF kernel train set are close in value. Similarly, the accuracy score of the linear SVM test set and the accuracy score of the linear SVM are close in value.

1.2

B. Average accuracy on the validation across the 10 folds for different C values

C values	Average Accuracy
C = 0.01	0.8141644079393544
C = 0.1	0.814594307931735
C = 1	0.81453288287031
C = 10	0.8145635859713233
C = 100	0.8145635954010226

C.

Highest accuracy score is: 0.814594307931735 when C = 0.1

Results when C = 0.1 is used to train a linear SVM on the whole training set X and tested on test data.

Accuracy: 0.8132178613107303

Confusion matrix: $\begin{bmatrix} 12087 & 348 \\ 2693 & 1153 \end{bmatrix}$

The accuracy and confusion matrix results for when C = 0.1 is used to train a linear SVM are nearly the same to the results from 1.1a.

1.3

B.

	C = 0.01	C = 0.1	C = 1	C = 10	C = 100
$10^{-2}\gamma_0$	0.75919068 47395027	0.76404312 31465927	0.803108302 3599143	0.815300733 8569205	0.829121042 6380828
$10^{-1}\gamma_0$	0.76404312 31465927	0.80289332 40746259	0.828752511 1289312	0.842634188 3933226	0.846626741 948357
γ_0	0.80166479 45570267	0.84023874 30135357	84705672680 80318	0.849359883 7204918	0.846380796 5304742
$10^1\gamma_0$	0.78179439 06867763	0.82024530 04264486	0.843125456 8689326	0.830165306 4011063	0.814103171 4719158

$10^2\gamma_0$	0.75919068 47395027	0.763766766 9483758	0.790362988 7318865	0.779153028 1782046	0.774515615 2048998
----------------	------------------------	------------------------	------------------------	------------------------	------------------------

C.

Highest accuracy during the cross validation: $C = 10$ and $\gamma = \gamma_0$ (0.035399836257649225)

Accuracy: 0.8512376389656655

Confusion matrix: $\begin{bmatrix} 11720 & 715 \\ 1707 & 2139 \end{bmatrix}$

The results for when the parameters of the RBF kernel are: $C = 10$ and $\gamma = \gamma_0$ is close to the results of 1.1b, but it has a slightly higher accuracy score. The results for when the parameters of the RBF kernel are: $C = 10$ and $\gamma = \gamma_0$ is higher than the results from 1.2c as the accuracy score of this RBF kernel is 0.8512376389656655 and the accuracy score of 1.2c is 0.8132178613107303. As a result, the results for this RBF kernel is higher than the results from 1.2c.

2.1

Train Set:

Accuracy: 0.9045483861060778

Confusion matrix: $\begin{bmatrix} 23672 & 1048 \\ 2060 & 5781 \end{bmatrix}$

Test Set:

Accuracy: 0.8696640255512561

Confusion matrix: $\begin{bmatrix} 11663 & 772 \\ 1350 & 2496 \end{bmatrix}$

2.2

Default Parameters:

```
{'objective': 'binary:logistic', 'base_score': 0.5, 'booster': 'gbtree', 'colsample_bylevel': 1, 'colsample_bynode': 1, 'colsample_bytree': 1, 'gamma': 0, 'gpu_id': -1, 'importance_type': 'gain', 'interaction_constraints': '', 'learning_rate': 0.300000012, 'max_delta_step': 0, 'max_depth': 6, 'min_child_weight': 1, 'missing': nan, 'monotone_constraints': '()', 'n_estimators': 100, 'n_jobs': 0, 'num_parallel_tree': 1, 'random_state': 0, 'reg_alpha': 0, 'reg_lambda': 1, 'scale_pos_weight': 1, 'subsample': 1, 'tree_method': 'exact', 'validate_parameters': 1, 'verbosity': None}
```

Parameters that changed

1. Parameter changed: learning rate

Default learning rate: 0.300000012
Average accuracy score: 0.8709808735522584

Learning rate: 0.05
Average Accuracy Score: 0.8673876394746827

Learning rate: 0.1
Average Accuracy Score: 0.8724549712997671

Learning rate: 0.15
Average Accuracy Score: 0.872332177755113

Learning rate: 0.2
Average Accuracy Score: 0.8717792767646928

Learning rate: 0.25
Average Accuracy Score: 0.8715949732913195

2. Parameter changed: max_depth

Default max_depth : 6
Average Accuracy Score: 0.8709808735522584

Max_depth: 3
Average Accuracy Score: 0.8712879705702855

Max_depth: 4
Average Accuracy Score: 0.8722093182025634

Max_depth: 5
Average Accuracy Score: 0.8717486585309736

Max_depth: 7

Average Accuracy Score: 0.8691073997490946

Max_depth: 8

Average Accuracy Score: 0.8675411266906508

3. Parameter changed: reg_lambda

Default reg_lambda : 1

Average Accuracy Score: 0.8709808735522584

reg_lambda: 0.01

Average Accuracy Score: 0.8712265832276579

reg_lambda: 0.1

Average Accuracy Score: 0.869752315745561

reg_lambda: 10

Average Accuracy Score: 0.8709193447641406

reg_lambda: 0.5

Average Accuracy Score: 0.8693530622759974

reg_lambda: 0

Average Accuracy Score: 0.8694451621493376

4. Parameter Changed: min_child_weight

Default min_child_weight : 1

Average Accuracy Score: 0.8709808735522584

Min_child_weight: 0.1

Average Accuracy Score: 0.8695988190998938

Min_child_weight: 0.01

Average Accuracy Score: 0.8701208849735101

Min_child_weight: 5

Average Accuracy Score: 0.869598687084103

Min_child_weight: 10

Average Accuracy Score: 0.8707658952669698

Min_child_weight: 50

Average Accuracy Score: 0.8675411172609516

5. Parameter Changed: gamma

Default gamma : 0

Average Accuracy Score: 0.8709808735522584

Gamma: 0.5

Average Accuracy Score: 0.8721785868124524

Gamma: 1

Average Accuracy Score: 0.8709808735522584

Gamma: 0.25

Average Accuracy Score: 0.8731920248883712

Gamma: 0.75

Average Accuracy Score: 0.8719635048004714

Gamma: 0.8

Average Accuracy Score: 0.8700595164902809

6. Parameter Changed: colsample_bytree

Default colsample_bytree : 1

Average Accuracy Score: 0.8709808735522584

colsample_bytree : 0

Average Accuracy Score: 0.8699365343516403

colsample_bytree : 0.5

Average Accuracy Score: 0.8721785868124524

colsample_bytree : 0.25

Average Accuracy Score: 0.8731920248883712

colsample_bytree : 0.75

Average Accuracy Score: 0.8719635048004714

colsample_bytree : 0.8

Average Accuracy Score: 0.8700595164902809

7. Parameter Changed: reg_alpha

Default reg_alpha : 0

Average Accuracy Score: 0.8709808735522584

reg_alpha : 0.5
Average Accuracy Score: 0.8702744381973735

Reg_alpha: 1
Average Accuracy Score: 0.869875071571418

Reg_alpha: 1.5
Average Accuracy Score: 0.8693530811353961

reg_alpha : 2
Average Accuracy Score: 0.8693530811353961

Reg_alpha: 2.5
Average Accuracy Score: 0.8693530999947947

The parameters with the highest accuracy scores are:

Learning_rate = 0.1
max_depth = 4
Reg_lambda = 0.01
Min_child_weight = 1
Gamma = 0.25
Colsample_bytree: 0.25
Reg_alpha = 0

Accuracy: 0.8697868681284934
Confusion matrix: $\begin{bmatrix} 11822 & 613 \\ 1507 & 2339 \end{bmatrix}$