Implementation Instruction

[1] Download QVT.R[1], then open it using free software R[2] or R Studio[3].

[2] Lines 3-7: show the required R packages. If these packages have been downloaded before, run lines 3-7 directly, or install the missing packages using **install.packages("XX")** function.

[3] Lines 10-29: the function of computing macro-F, which works as the performance metric. The other performance metric AUC will be obtained using package ROCR.

[4] Line 32: read in the data. Here the read.delim() function is used as an example.

[5] Lines 35-44: the data preprocessing, including making the small-sample-size data. Here 50 samples are extracted from each class randomly. The value can be changed.

[6] Lines 47-52: make the training data and testing data. The training data are used to build learning model, and the testing data are used to evaluate the performance of learning model.

[7] Lines 55-57: perform the data normalization.

[8] Lines 60-168: the implementation of QVT.

[9] Line 170: output the result of feature selection (data: halloffame, feature selection method base: lasso), which is shown as follows. The **num** is the number of column of variable in the original data, **score** is the importance score.

```
> ig_sort1
                num        score
On.base.pct      12 2.4801340512
Fielding.ave     14 2.2942673655
Batting.average  11 1.4482220093
diff.vector      13 0.4981327374
Doubles           5 0.0008882277
Hits              4 0.0001709897
RBIs              8 0.0001074237
```

[10] Lines 173-210: using classifier svm to build the learning model, and test the performance. The **numOfvari** is the used number of top features in building the learning model. **QVT(Lasso)_SVM_F** is the value of macro_F, **QVT(Lasso)_SVM_AUC** is the value of AUC.

```
> results_of_learningModel(ig_sort1, Train, Test)
     numOfvari QVT(Lasso)_SVM_F QVT(Lasso)_SVM_AUC
[1,]         2        0.8000000          0.8304498
[2,]         3        0.8108108          0.8408304
[3,]         4        0.8333333          0.8546713
[4,]         5        0.7777777          0.8927336
[5,]         6        0.7777777          0.9307958
[6,]         7        0.8000000          0.9273356
```

[1] https://github.com/zhengwanw/QVT/blob/a87d976be5216b9bfe6623e85aeb4ec3545eab3f/QVT.R

[2] https://cran.r-project.org/

[3] https://www.rstudio.com/products/rstudio/download/