**Name: Chaofan Hu**

Github URL: https://github.com/ChaofanHu/MobileComputingAssignment04.git

**EXERCISE 0**

**5-FOLD CROSS VALIDATION**

Average validation score of XGBoost: 63.40 ± 0.67 %

Average validation score of DummyClassifier: 50.49 ± 0.24 %

**LEAVE-ONE-USER-OUT CROSS VALIDATION**

Average validation score of XGBoost: 41.92 ± 16.32 %

Average validation score of DummyClassifier: 49.83 ± 1.21 %

Compared with the former data, the result has changed.

**EXERCISE 1**

**A screenshot of a computer program

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

**LEAVE-ONE-DAY-OUT CROSS VALIDATION**

Average validation score of XGBoost: 64.05 ± 1.03 %

Average validation score of DummyClassifier: 49.78 ± 0.52 %

This result is better than in the other validation paradigms.

Leave-One-User-Out method, the model's lower performance and higher variability indicate sensitivity to user-specific characteristics, possibly due to overfitting to unique patterns in user data that do not generalize well.

Leave-One-Day-Out approach shows better performance, suggesting that day-to-day data variability is less pronounced compared to user-to-user variability, allowing the model to find more consistent and generalizable patterns.

**EXERCISE 2**

A screenshot of a computer program

Description automatically generated

T-Test for 5-fold

In this result, There is a significant difference in the performance of the two models.

A screenshot of a computer program

Description automatically generated

T-test for leave one user

In this result, There is NO significant difference in the performance of the two models.

A screenshot of a computer program

Description automatically generated

T- test for Leave One Day

In this result, there is a significant difference in the performance of the two models.

**EXERCISE 3**

1. 5-fold cross validation

A screen shot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated A screenshot of a computer program

Description automatically generated

t-test for svc and dummy

t-statistic: -3.3715766721643154, p-value: 0.009762740154260326

There is a significant difference in the performance of the two models.

t-test for rf and dummy

t-statistic: 30.65531409883557, p-value: 1.3928926522419737e-09

There is a significant difference in the performance of the two models.

t-test for nb and dummy

t-statistic: 7.853078668192827, p-value: 4.988910262028665e-05

There is a significant difference in the performance of the two models.

t-test for svc and xgboost

t-statistic: -39.20592384520641, p-value: 1.969192019630918e-10

There is a significant difference in the performance of the two models.

t-test for rf and xgboost

t-statistic: 0.4889765974928941, p-value: 0.6379813066067985

There is no significant difference in the performance of the two models.

t-test for nb and xgboost

t-statistic: -13.57516088340624, p-value: 8.332877276027414e-07

There is a significant difference in the performance of the two models.

2. Leave One User Out Cross Validation

A screenshot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated

t-test for svc and dummy

t-statistic: 0.37446341455249504, p-value: 0.7129761246537354

There is no significant difference in the performance of the two models.

t-test for rf and dummy

t-statistic: -1.7102207114376262, p-value: 0.10012398132663432

There is no significant difference in the performance of the two models.

t-test for nb and dummy

t-statistic: -3.2918750351918753, p-value: 0.0030721976345413674

There is a significant difference in the performance of the two models.

t-test for svc and xgboost

t-statistic: 1.050002404359758, p-value: 0.3093187819010649

There is no significant difference in the performance of the two models.

t-test for rf and xgboost

t-statistic: -0.011775956635890273, p-value: 0.9907017073883206

There is no significant difference in the performance of the two models.

t-test for nb and xgboost

t-statistic: -1.6809783240036107, p-value: 0.10573743863615517

There is no significant difference in the performance of the two models.

3. Leave One Day Out Cross Validation

The codes are similar to Leave One User Validation, the result is:

t-test for svc and dummy

t-statistic: 1.047140954404872, p-value: 0.3174986035477387

There is no significant difference in the performance of the two models.

t-test for rf and dummy

t-statistic: 21.025066411217548, p-value: 5.455418666449534e-12

There is a significant difference in the performance of the two models.

t-test for nb and dummy

t-statistic: 8.742963357154382, p-value: 4.799396284128512e-07

There is a significant difference in the performance of the two models.

t-test for svc and xgboost

t-statistic: -27.739655211227245, p-value: 1.5618403400253588e-11

There is a significant difference in the performance of the two models.

t-test for rf and xgboost

t-statistic: 0.050564539405819155, p-value: 0.9603868684628336

There is no significant difference in the performance of the two models.

t-test for nb and xgboost

t-statistic: -14.028601735962079, p-value: 1.2300288988691661e-09

There is a significant difference in the performance of the two models.