GridSearchForParams

April 9, 2022

[1]:
```
import pandas as pd
import numpy as np
from surprise import Reader
from surprise import KNNBasic
from surprise import Dataset
from surprise import KNNWithMeans
from surprise import SVD
import timeit
from surprise.model_selection import GridSearchCV
from surprise.model_selection import cross_validate

reader = Reader()
data_df = pd.read_csv('ratings_u.data',sep='\t',names=['uid','iid','rating'])
whole_data = Dataset.load_from_df(data_df, reader=reader)
data_df
SVD_param_grid = {
    'n_factors': [5,10,20,40,80],
    'n_epochs': [5,10,15],
    'biased': [True],
}
SVD_start=timeit.default_timer()
gs_cv = 3
# eval_cv = 5
SVD_gs = GridSearchCV(SVD, SVD_param_grid, measures=['rmse'], cv=gs_cv)
SVD_gs.fit(whole_data)
SVD_based_algo = SVD_gs.best_estimator['rmse']

SVD_best_result = cross_validate(SVD_based_algo, whole_data, measures=['RMSE'],␣
↪verbose=True)

SVD_mean_rmse = SVD_best_result['test_rmse'].mean()
SVD_runtime=timeit.default_timer()-SVD_start
print(SVD_gs.best_params,SVD_mean_rmse,SVD_runtime)

KNN_param_grid = {'k': [5,15,35,75,115],
                  'sim_options': {
                  'name': ['cosine', 'pearson'],
```

```
                        'user_based': [False,True]
                    }
                    }

KNN_start=timeit.default_timer()
KNN_gs = GridSearchCV(KNNWithMeans, KNN_param_grid, measures=['rmse'], cv=gs_cv)
KNN_gs.fit(whole_data)

KNN_algo = KNN_gs.best_estimator['rmse']

KNN_best_result = cross_validate(KNN_algo, whole_data, measures=['rmse'],␣
↪verbose=True)

KNN_mean_rmse = KNN_best_result['test_rmse'].mean()
KNN_runtime=timeit.default_timer()-KNN_start
print(KNN_gs.best_params, KNN_mean_rmse,KNN_runtime)
```

```
Evaluating RMSE of algorithm SVD on 5 split(s).

                  Fold 1  Fold 2  Fold 3  Fold 4  Fold 5  Mean    Std
RMSE (testset)    0.8802  0.8805  0.8778  0.8802  0.8765  0.8790  0.0016
Fit time          13.44   13.48   13.45   13.43   13.52   13.46   0.03
Test time         1.35    1.18    1.49    1.35    1.34    1.34    0.10
{'rmse': {'n_factors': 40, 'n_epochs': 15, 'biased': True}} 0.879024130314346
494.23412910000013
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
```

```
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
```

```
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
```

```
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Evaluating RMSE of algorithm KNNWithMeans on 5 split(s).

                  Fold 1  Fold 2  Fold 3  Fold 4  Fold 5  Mean    Std
RMSE (testset)    0.8848  0.8850  0.8847  0.8874  0.8848  0.8853  0.0010
Fit time          29.28   29.57   29.50   29.44   29.39   29.44   0.10
Test time         56.86   55.39   56.22   55.47   56.32   56.05   0.55
{'rmse': {'k': 115, 'sim_options': {'name': 'pearson', 'user_based': False}}}
0.8853432813538428 8020.885591199999
```

[2]:
```
SVD_algo = SVD(n_factors=40, n_epochs=5, biased=True)

SVD_result = cross_validate(SVD_algo, whole_data, measures=['rmse'],␣
 ↪cv=5,verbose=True)
```

```
Evaluating RMSE of algorithm SVD on 5 split(s).

                  Fold 1  Fold 2  Fold 3  Fold 4  Fold 5  Mean    Std
RMSE (testset)    0.9116  0.9118  0.9122  0.9137  0.9129  0.9124  0.0008
Fit time          4.60    4.73    4.75    4.79    4.81    4.73    0.07
Test time         1.68    1.67    1.68    1.68    1.42    1.63    0.10
```

[4]:
```
start=timeit.default_timer()
KNN_algo = KNNWithMeans(k=15, sim_options={
      "name": "pearson",
      "user_based": False,
    })

KNN_result = cross_validate(KNN_algo, whole_data, measures=['rmse'],␣
 ↪cv=5,verbose=True)
runtime=timeit.default_timer()-start
print(runtime)
```

```
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
```

```
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Evaluating RMSE of algorithm KNNWithMeans on 5 split(s).

                  Fold 1  Fold 2  Fold 3  Fold 4  Fold 5  Mean    Std
RMSE (testset)    0.9074  0.9052  0.9086  0.9075  0.9074  0.9073  0.0011
Fit time          29.26   29.27   29.32   29.38   29.51   29.35   0.09
Test time         36.32   35.78   36.60   36.61   36.11   36.28   0.31
333.5900452999995
```

[5]:
```
from scipy.stats import ttest_ind

SVD_RMSE1 = SVD_best_result['test_rmse']
SVD_RMSE2 = SVD_result['test_rmse']

ttest_ind(SVD_RMSE1,SVD_RMSE2)
```

[5]:
```
Ttest_indResult(statistic=-38.10367009609319, pvalue=2.471109834820134e-10)
```

[6]:
```
KNN_RMSE1 = KNN_best_result['test_rmse']
KNN_RMSE2 = KNN_result['test_rmse']

ttest_ind(KNN_RMSE1,KNN_RMSE2)
```

[6]:
```
Ttest_indResult(statistic=-29.07475239259398, pvalue=2.1201287020547553e-09)
```

[7]:
```
gs_cv = 3
KNN_param_grid = {'k': [5,15,25],
                    'sim_options': {
                    'name': ['cosine', 'pearson'],
                        'user_based': [False,True]
                    }
                    }

KNN_start=timeit.default_timer()
KNN_gs2 = GridSearchCV(KNNWithMeans, KNN_param_grid, measures=['rmse'],␣
 ↪cv=gs_cv)
KNN_gs2.fit(whole_data)

KNN_algo2 = KNN_gs2.best_estimator['rmse']

KNN_best_result2 = cross_validate(KNN_algo2, whole_data, measures=['rmse'],␣
 ↪verbose=True)

KNN_mean_rmse2 = KNN_best_result2['test_rmse'].mean()
KNN_runtime=timeit.default_timer()-KNN_start
```

```
print(KNN_gs2.best_params, KNN_mean_rmse2,KNN_runtime)
```

```
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
```

```
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Evaluating RMSE of algorithm KNNWithMeans on 5 split(s).

                  Fold 1  Fold 2  Fold 3  Fold 4  Fold 5  Mean    Std
RMSE (testset)    0.8940  0.8938  0.8972  0.8964  0.8933  0.8949  0.0016
Fit time          29.35   29.88   29.41   29.61   30.04   29.66   0.27
Test time         39.61   39.69   39.44   39.34   40.16   39.65   0.28
{'rmse': {'k': 25, 'sim_options': {'name': 'pearson', 'user_based': False}}}
0.8949207942484836 4749.8469742
```

[2]:
```
import pandas as pd
import numpy as np
```

```
from surprise import Reader
from surprise import KNNBasic
from surprise import Dataset
from surprise import KNNWithMeans
from surprise import SVD
import timeit
from surprise.model_selection import GridSearchCV
from surprise.model_selection import cross_validate

reader = Reader()
data_df = pd.read_csv('ratings_u.data',sep='\t',names=['uid','iid','rating'])
whole_data = Dataset.load_from_df(data_df, reader=reader)

KNN_param_grid = {'k': [115,215,315],
                    'sim_options': {
                    'name': ['cosine', 'pearson'],
                        'user_based': [False,True]
                    }
                    }

KNN_start=timeit.default_timer()
KNN_gs = GridSearchCV(KNNWithMeans, KNN_param_grid, measures=['rmse'], cv=3)
KNN_gs.fit(whole_data)

KNN_algo = KNN_gs.best_estimator['rmse']

KNN_best_result = cross_validate(KNN_algo, whole_data, measures=['rmse'],␣
 ↪verbose=True)

KNN_mean_rmse = KNN_best_result['test_rmse'].mean()
KNN_runtime=timeit.default_timer()-KNN_start
print(KNN_gs.best_params, KNN_mean_rmse,KNN_runtime)
```

```
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
```

```
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the cosine similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
```

```
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Computing the pearson similarity matrix…
Done computing similarity matrix.
Evaluating RMSE of algorithm KNNWithMeans on 5 split(s).

                  Fold 1  Fold 2  Fold 3  Fold 4  Fold 5  Mean    Std
RMSE (testset)    0.8861  0.8837  0.8842  0.8862  0.8860  0.8852  0.0011
Fit time          56.00   48.63   56.21   56.21   54.87   54.38   2.92
Test time         86.81   105.11  107.17  107.37  96.84   100.66  7.92
{'rmse': {'k': 115, 'sim_options': {'name': 'pearson', 'user_based': False}}}
0.8852391478833634 10273.794491
```