Code guide - Item-based ACF Recommendation

Rémi Pichon - 4 mars 2016

What has been implemented

Similarity metrics

Several similarity metrics have been implemented :

- CosineMetric (created) experiment 1 & 2

Implementation of the Cosine metric (see formula in report)

- JaccardMetric (created) experiment 3

Implementation of the Jaccard Index (see formula in report)

- PearsonSignifianceWeightMetric (created) experiment 3

From PearsonMetric, add Signifiance Weighting. Parameters N is given to the constructor by the different runners.

- PearsonJaccardSignifianceWeightMetric (created) experiment 3

From PearsonMetric, add Signifiance Weighting with Jaccard Index extension.

Predictors

Two Predictors have been implemented :

- DeviationPredictor (created) experiment 1 & 3

Implementation of Deviation from Mean approach.

- WeightedAveragePredictor (created) experiment 1 & 3

Implementation of Deviation from Mean approach. Some code has been added to output the standard deviation for every item similarity weight (see report, experiment 1)

Neighbourhood

One Neighbourhood have been implemented :

- SimilarityThresholdingNeighbourhood (created) experiment 2

Implementation of neighbourhood formation based on similarity with a threshold L as abs(wa,i) > L given to the constructor by the different runners

Helpers

Some helpers have been used :

- Statistics (copier from internet)

Class to perform standard deviation (experiment 1)

- Profile (updated) experiment 3

Add method `Set<Integer> getUnionIds(final Profile other)` to easily compute Jaccard Index (see JaccardMetric)

- ItemBasedCF (updated) experiment 2

Add method `averageNeighbourhoodSize` to compute and neighbourhood size (used in experiment 2)

Running instruction

Several runners have been implemented :

- ExecuteParams

In order to easily switch between interesting couple *predictor* + *metric*, this class contains methods to get a predictor and a metric just by calling a specific method name as *predictor_metric()*.

- ExecuteIB_Iterate_Exp_1

Outputs RMSE and Coverage for neighbourhood size from 5 to 100.

- ExecuteIB_Exp_2

Outputs RMSE and Coverage for neighbourhood formation by similarity with threshold set a value (0.8)

- ExecuteIB_Iterate_Exp_2

Outputs RMSE, Coverage and neighbourhood size for neighborhood formation by similarity with threshold from 0 to 1

- ExecuteIB_Exp_3

Outputs RMSE and Coverage with nearest neighbourhood formation with neighbourhood size of 70, deviation from mean predictor and for the following Similarities form : Cosine, Pearson, Pearson Signifiance Weighting (1, 5, 50, 100, 200) and Jaccard