South China University of Technology

The Experiment Report of Machine Learning

SCHOOL: SCHOOL OF SOFTWARE ENGINEERING

SUBJECT: SOFTWARE ENGINEERING

Author:
GuoBing Wu,ShiHong Chen and HaoYuan Yao

Supervisor:
Qingyao Wu

Student ID：
201530613030
201530611234
201530613436

Grade:
Undergraduate or Graduate

December 9, 2017

Recommender System Based on Matrix Decomposition

Abstract— In this experiment, we implement the matrix decomposition based recommender system and conduct the experiment on the Movielens-100k dataset. According to the result, we noticed the importance of proper parameter setting, especially the parameter k. After the experiment, we have a better comprehension on recommender system and SGD algorithm.

I. INTRODUCTION

In this experiencement,we will achieve a recommender system based on matrix decomposition.We have four learning goals. Firstly,explore the construction of recommended system. Secondly,understand the principle of matrix decomposition. Thirdly,be familiar to the use of gradient descent. Finally,construct a recommendation system under small-scale dataset, cultivate engineering ability.

II. METHODS AND THEORY

Models are learned from the underlying data rather than heuristics

Give a rating matrix $\mathbf{R} \in \mathbb{R}^{m \times n}$ with sparse ratings from m users to n items.

Assume rating matrix R can be factorized into the multiplication of two low-rank feature matrices $\mathbf{P} \in \mathbb{R}^{m \times k}$ and $\mathbf{Q} \in \mathbb{R}^{k \times n}$.

SGD is to minimize the following objective function:

$$\mathcal{L} = \sum_{u,i \in \Omega} (r_{u,i} - \mathbf{p}_u^\top \mathbf{q}_i)^2 + \lambda_p ||\mathbf{p}_u||^2 + \lambda_q ||\mathbf{q}_i||^2$$

General Steps of SGD:

Require feature matrices P, Q, observed set Ω , regularization parameters λ p, λ q and learning rate α .

2: Randomly select an observed sample ru,i from observed set Ω .

3: Calculate the gradient w.r.t to the objective function.

$$\frac{\partial \mathcal{L}}{\partial \mathbf{p}_u} = E_{u,i}(-\mathbf{q}_i) + \lambda_p \mathbf{p}_u$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} = E_{u,i}(-\mathbf{p}_u) + \lambda_q \mathbf{q}_i$$

4: Update the feature matrices P and Q with learning rate α and gradient.

$$\mathbf{p}_u = \mathbf{p}_u + \alpha(E_{u,i}\mathbf{q}_i - \lambda_p \mathbf{p}_u)$$

$$\mathbf{q}_i = \mathbf{q}_i + \alpha(E_{u,i}\mathbf{p}_u - \lambda_q \mathbf{q}_i)$$

5: Repeat the above processes until convergence

III. EXPERIMENT

A. Dataset

We implement the experiment on Movielens-100k dataset. It consists 10000 comments from 943 users out of 1682 movies. We use the u1.data and u1.test for the splitting of training set and test set.

B. Result

1.Read the data set and divide it .Populate the original scoring matrix R_{n_users,n_items} against the raw data, and fill 0 for null values.

2.Initialize the user factor matrix $P_{n\_users,K}$ and the item (movie) factor matrix $Q_{n\_item,K}$, where K is the number of potential features.

3.Determine the loss function and hyperparameter learning rate η and the penalty factor λ.

4.Use the stochastic gradient descent method to decompose the sparse user score matrix, get the user factor matrix and item (movie) factor matrix:

4.1 Select a sample from scoring matrix randomly;

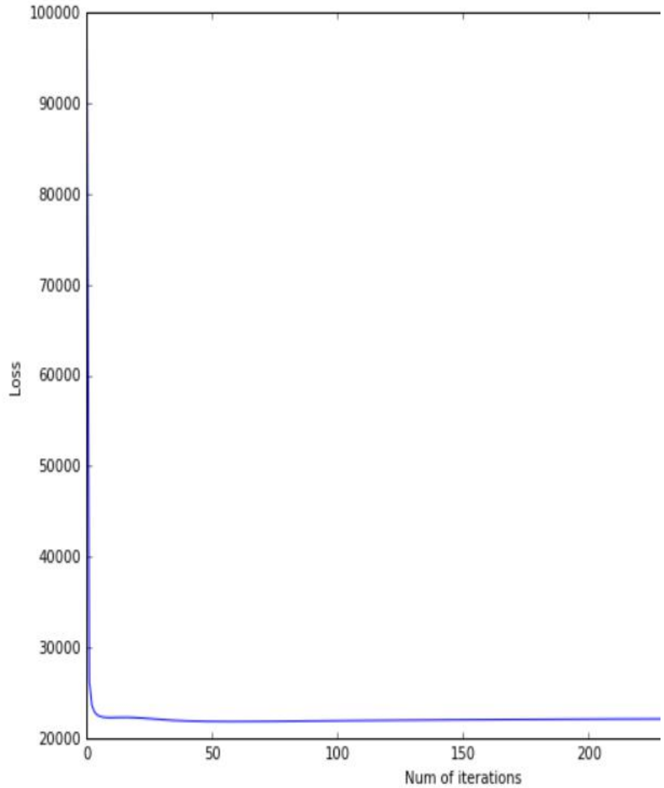
4.2 Calculate this sample's loss gradient of specific row(column) of user factor matrix and item factor matrix;

4.3 Use SGD to update the specific row(column) of $P_{n\_users,K}$ and $Q_{n\_item,K}$;

4.4 Calculate the on the validation set, comparing with the of the previous iteration to determine if it has converged.

5.Repeat step 4. several times, get a satisfactory user factor matrix P and an item factor matrix Q , Draw a $L_{validation}$ curve with varying iterations.

6.The final score prediction matrix $\hat{R}_{n\_users,n\_items}$ is obtained by multiplying the user factor matrix $P_{n\_users,K}$ and the transpose of the item factor matrix $Q_{n\_item,K}$.



IV. CONCLUSION

In this experiencement,we achieve a recommender system based on matrix decomposition.
We have learned a lot from this experiencement.
Firstly,explore the construction of recommended system.
Secondly,understand the principle of matrix decomposition.
Thirdly,be familiar to the use of gradient descent.
Finally,construct a recommendation system under small-scale dataset, cultivate engineering ability.