South China University of Technology

The Experiment Report of Machine Learning

SCHOOL: SCHOOL OF SOFTWARE ENGINEERING

SUBJECT: SOFTWARE ENGINEERING

Author:

饶浩聪. 闫明远 and 谢家峻

Supervisor:

Qingyao Wu

Student ID:

201530612644
201530613320
201530613191

Grade:

Undergraduate

December 9, 2017

Recommender System Based on Matrix Decomposition

Abstract—

I. INTRODUCTION

Motivation of Experiment:

1. Explore the building process of recommendation system
2. Understand the principle of matrix decomposition
3. Skillful use of gradient descent
4. Implement recommendation system on simple small data set, and train engineering capability

In recent years, recommendation system has been widely used. The system recommend items for its latent customer based on the previous purchase record that the customer or the similar customer had.

However, we haven't realized a recommendation system in machine learning class. In order to explore the construction of recommended system, understand the principle of matrix decomposition, getting us familiar to the use of SGD, and construct a recommendation system under small-scale dataset, we conduct the recommender system experiment.

In this experiment, we implement the matrix decomposition based recommender system on MovieLens-100k dataset, achieving a satisfying result.

II. METHODS AND THEORY

Recommender System applies statistical and knowledge discovery techniques to the problem of making product recommendations.

The Matrix Factorization is a model based collaborative filtering algorithm. It is based on a rating matrix, with sparse ratings form m users to n items.

It assume rating matrix R can be factorized into the multiplication of two low-rank feature matrices P and Q. We use SGD to get the proper P and Q in order to fill the matrix. The objective function of SGD is in figure 1, where r denotes the actual rating of user u for item I, Ω denotes the set of observed samples from rating matrix R, λ p and λq are regularization parameters to avoid overtting.

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

Figure 1 The objective function of SGD

In order to apply SGD to obtain the proper P and Q, we calculate the gradient of P and Q. The formula is in figure 2.

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

Figure 2 The gradient of P and Q

Algorithm 1 SGD

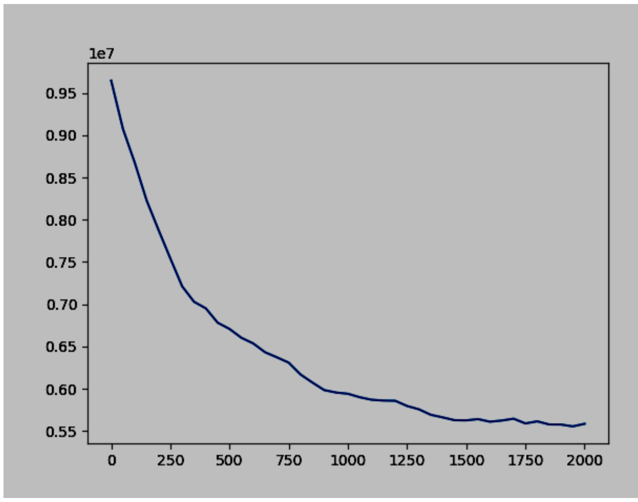
1: **Require** feature matrices $\mathbf{P}$, $\mathbf{Q}$, observed set Ω, regularization parameters λ_p, λ_q and learning rate α.
2: **Randomly** select an observed sample $r_{u,i}$ from observed set Ω.
3: Calculate the **gradient** w.r.t to the objective function:
$$E_{u,i} = r_{u,i} - \mathbf{p}_u^\top \mathbf{q}_i$$
$$\frac{\partial \mathcal{L}}{\partial \mathbf{p}_u} = E_{u,i}(-\mathbf{q}_i) + \lambda_p \mathbf{p}_u$$
$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} = E_{u,i}(-\mathbf{p}_u) + \lambda_q \mathbf{q}_i$$
4: **Update** the feature matrices $\mathbf{P}$ and $\mathbf{Q}$ 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

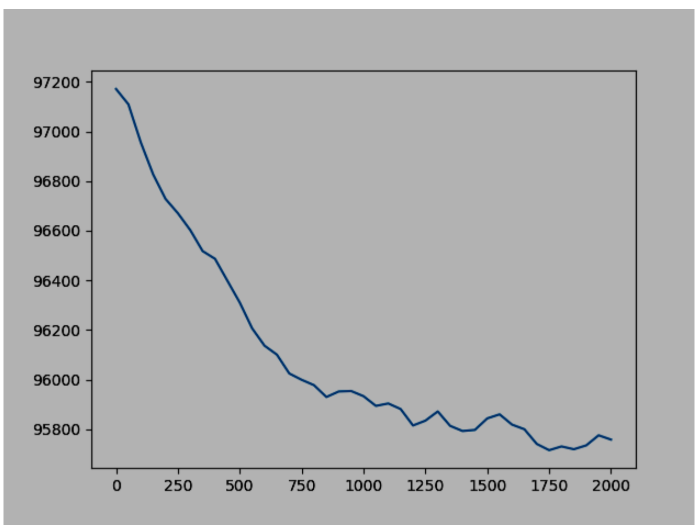
We implement the matrix decomposition based recommender system using python3 and anaconda toolkit. We set the regularization parameters to 0.3, set the learning Rate to 0.01, set K to 100 and make 2000 iterations. According to the experiment result, we can see that during the iterations, the loss value drops significantly.

We also noticed that different k value may lead to different result. In this experiment, we set K to 10, the loss value drops relatively smaller than when k equals to 100. However, when we set k to 1000, the experiment failed.

We also realized that the result also rely on the initial value of P and Q matrix. If the initial value is too large, there may be an overflow value during the experiment. So we set the initial value to a value that between 0 and 1.



experiment result with k = 100



experiment result with k = 10

Code content:

Recommendation.py

IV. CONCLUSION

In this experiment, we implement the recommender system based on matrix decomposition. After this experiment, we have got a deeper comprehension to recommender system.

We also get a deeper knowledge on SGD algorithm and its application.

We realize the important of setting parameters one more time. Meanwhile, we realize the role of parameter k in this algorithm.

After the experiment, we are more interested in recommender system, and we will read more paper about this area.