华南理工大学

South China University of Technology

The Experiment Report of Machine Learning

SCHOOL: SCHOOL OF SOFTWARE ENGINEERING

SUBJECT: SOFTWARE ENGINEERING

Author: KongJunSheng，ZhangYuan and ShenChengSheng

Supervisor: Mingkui Tan

Student ID: 201530611852, 201530613696 , 201530612682

Grade: Undergraduate

December 9, 2017

Linear Regression, Linear Classification and Gradient Descent

Abstract—Recommendation system based on matrix decomposition

I. INTRODUCTION

1 to explore recommendation system building process
2 to understand the principle of matrix decomposition
3 proficient use of gradient descent
4 to realize recommendation system in a simple small data sets, cultivating engineering ability

II. METHODS AND THEORY

1.Using the stochastic gradient descent method to optimize
2.The Loss function is used by the square of error between the original score matrix R(m*n) and the rebuilt score matrix R'(m*n)

III. EXPERIMENT

A. Dataset

Dataset u1.base / u1.test-making up of score that 943 users article 1682 of the 10000 movie.Each user to score at least 20 films.Users and film are numbered consecutively from 1.The data is random sequence.

B. Implementation

1. Read the data, and divide the data set (or directly through u1. Base/u1. Test to u5. The base/u5 test).According to the original data fill the original score matrix R (n_users n_items), for a null value can fill is zero.
2. Initialize the user factor matrix and items (film) factor matrix, which is a potential feature Numbers.
3. Determine the loss function and determine the super parameter vector and penalty coefficient.
4. The stochastic gradient descent method is used to decompose the sparse users rating matrix, get the user factor matrix (film) factor matrix and items:
   - randomly selected user rating matrix of a sample;
   - to calculate the cost of this sample function value for the user factor matrix a row (column) and item factor matrix a row (column) of gradient;
   - update user of gradient descent factor matrix row (column) and item factor matrix row (column);
   - calculation on the validation set loss function, can be compared with the previous iteration of loss function to judge whether the convergence.
5. Repeat step 4. Several times, to get satisfactory user factor matrix and items factor matrix, draw the curve of loss function change with the number of iterations.
6. The user factor is multiplied by factor matrix transpose matrix and items can get the final score prediction matrix.

IV. CONCLUSION