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:
Lizhao Liu

Supervisor:
Mingkui Tan

Student ID:
201730683109

Grade:
Undergraduate

November 8, 2019

Linear Regression, Ridge Regression and Gradient Descent

***Abstract*—In this report, we solve the Housing Dataset by using Linear Regression, Ridge Regression by closed form solution, and even optimize Linear Regression by Using Gradient Descen and its variants e.g. Stochastic Gradient Descent and Mini Batch Gradient Descent.**
We perform experiments on four aspects:
1. Different lambdas in Ridge Regression
2. Comparision between Gradient Descent and its variants in terms of convergence and time-cost.
3. Tuning learning rate of Mini-Batch Gradient Descent.
4. Comparing the magnitude of W between linear regression optimized by closed form solution, ridge regression and linear regression optimized by mini-batch Gradient Descent.

I. Introduction

THIS section introduces the problem to solved and leads the reader on to the main part. Detailed motivation is necessary. What's more, you can show your expected results and contributions.

II. Methods and Theory

In this section, you are asked to give a complete introduction to the experiment. For instance, the chosen methods, the related theories, the related equations(loss function), the derivation process(taking the gradient) and so on.

III. Experiments

A. Dataset

This section represents the related information of datasets, such as the content, the number of data, the training set, the validation set and so on.

B. Implementation

All detailed implementation in your experiment: initialization, process, results, all kinds of parameters. In a word, describe clearly What you do and how you do.

Figures and tables should be labeled and numbered, such as in Table I and Fig. 1.

TABLE I
Simulation Parameters

Information message length	$k = 16000$ bit
Radio segment size	$b = 160$ bit
Rate of component codes	$R_{cc} = 1/3$
Polynomial of component encoders	$[1, 33/37, 25/37]_8$

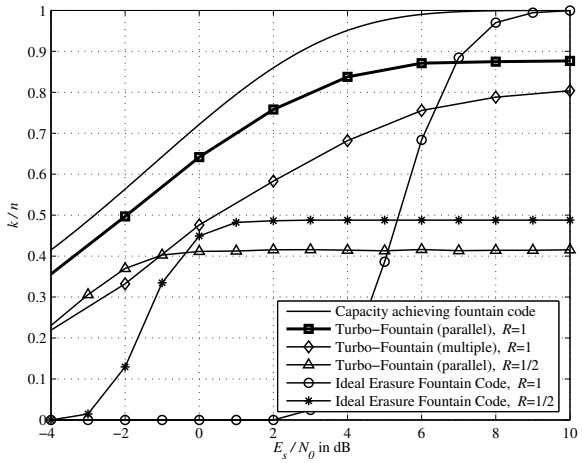

Fig. 1. Simulation results on the AWGN channel. Average throughput k/n vs E_s/N_0.

IV. Conclusion

This section summarizes the paper. In our experiments, you can also write your gains and inspirations in here.