

To create a title - feel free

Project report of Name

May 5, 2017

- Student:
Dang Vu Lam USTH BI4-078
Information and Communication
Technology Department
University of Science and Technology
of Hanoi
Hanoi, Vietnam
lam.dv@live.com
- Supervisors:
Sebastián Basterrech
and

Abstract

To be done later, it is a short text describing the thesis, around 500 characters (between 5 and 10 sentences).

Contents

1	Introduction	4
1.1	Specification of the notation	4
2	State of the art	4
2.1	Supervised Learning	4
2.2	Linear Regression Models	4
2.2.1	Model Form	4
2.2.2	Normal Equation	5
2.2.3	Geometric Interpretation of the solution	5
2.2.4	Computation of the Prediction of a Linear Model	5
3	Description of Linear Model	5
3.1	Extreme Learning Machines	5
3.2	Echo State Networks	6
3.3	Progressive Nets	6
4	Experimental Results	6
5	Conclusions	6

1 Introduction

The introduction should contain:

- Why is interesting to study the subject of your thesis, in what we can use our discover, etc.

(around 4-5 sentences)

- Description of the contributions. (around 4-5 sentences)

All the introduction should be around one page and half, or two pages, we will well define later.

1.1 Specification of the notation

To discuss together.

2 State of the art

2.1 Supervised Learning

Supervised Learning is a Statistical Learning technique in which for each observation of the predictor measurement x_i there is an associated response measurement y_i . [1, p. 26] The target is to fit a model that related the relation between the predictor and the response for prediction or inference.

2.2 Linear Regression Models

2.2.1 Model Form

$$Y = F(X) + \varepsilon = \beta_0 + \beta_1 * X_1 + \dots + \beta_n * X_n + \varepsilon$$

where Y is the response, X is the predictor. The Linear Model assume the linear relationship between the predictor X and response Y. [1]

2.2.2 Normal Equation

In order to solve the weights for linear regression problem, we compute the least square line/plane. The least square vector is the solution of the normal equation:

$$(X^T X)\hat{\beta} = X^T Y$$

One popular method of calculating a model $\hat{\beta}$ is to calculate the *Moore - Penrose generalized inverse* of X , denoted X^t . $\hat{\beta}$ then given as

$$\hat{\beta} = X^t Y [2]$$

2.2.3 Geometric Interpretation of the solution

2.2.4 Computation of the Prediction of a Linear Model

After the weight for a linear system have been calculated, prediction \hat{Y} can be obtained by

$$\hat{Y} = \hat{F}(X),$$

where \hat{F} represent the estimation of $F(X)$. The form of $\hat{F}(X)$ is not relevant as long as it yield accurate \hat{Y} [1, p. 19].

3 Description of Linear Model

3.1 Extreme Learning Machines

To be finished in the week 3 and week 4

Should be the model definition.

Also to insert some paragraph with the applications. Something like: this model has been used in computer vision [?]. Besides, in [] the authors applied ELM for

3.2 Echo State Networks

Model definition. Model properties. Also insert a paragraph about the applications.

3.3 Progressive Nets

Model definition. Model properties. Also insert a paragraph about the applications.

4 Experimental Results

We should select three problems of UCI dataset which are on the same field (for example you select applications in biology) and we will apply the three models and show and analyse the results.

If you are not sure about what topic, or dataset to select, we can talk about it and I can provide you dataset for analyzing.

At the beginning just include and analyze the results of:

- Salary dataset
- House prices

5 Conclusions

[To be done in the near future](#)

References

- [1] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning*, volume 103 of *Springer Texts in Statistics*. Springer New York. DOI: 10.1007/978-1-4614-7138-7.

- [2] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew. Extreme learning machine: Theory and applications. 70(1):489–501.