



UNSW

THE UNIVERSITY OF NEW SOUTH WALES

Student number: Z5085805

Student name: Sun YongXin

Report

Project

Introduction

This project teaches us how using the logistic regression classifier.

The goal of the project is to achieve the best classification accuracy by configuring your classifier and performing feature engineering.

Question

What feature do you use in your classifier?

Answer: Firstly, during design process, selection correct feature is important role with output result. I clearly understand the program cannot classification the characteristic of currently word. So, only one way, we should not only consider the word itself, but also the context. Therefore, I design different feature on my classifier. Taking many test, I find out the following feature, which has best result.

```
[[{ "ABC" : "Prime" , " DEFG" : " NNP" , " A" : "Minister" , " ATag" : " NNP" },  
{ "ABC" : " Minister" , " DEFG" : " NNP" , "B" : " Prime" , "BTag" : " NNP" , "A" : " Malcolm" ,  
" ATag" : " NNP" },  
{ "ABC" : " Malcolm" , " DEFG" : " NNP" , "B" : " Minister" , " BTag" : " NNP" , "A" : " Turnbull" ,  
" ATag" : " NNP" }...]]
```

I download data.dat from website, to test my program, and compare with result of F_r.