**Data Mining**

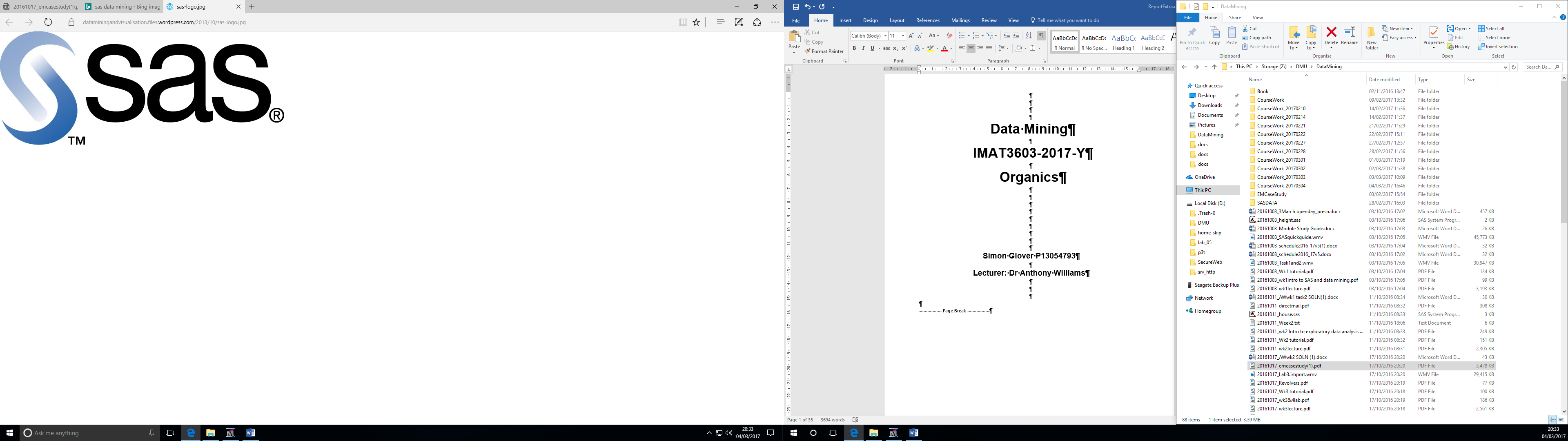
**IMAT5238**

**PNumber**

**Date**

**Framingham Assignment**

(Word count)



**Lecturer: Dr Anthony Williams**

**Table of contents**

Contents

[Summary 4](#_Toc67411278)

[Business Problem 4](#_Toc67411279)

[Data Mining Representation 4](#_Toc67411280)

[Methodology - data mining approach 4](#_Toc67411281)

[Data Exploration 4](#_Toc67411282)

[Data partition creation of model sets 4](#_Toc67411283)

[Data Modelling 5](#_Toc67411284)

[Regression 5](#_Toc67411285)

[Modification Node 5](#_Toc67411286)

[Selection Methods 5](#_Toc67411287)

[Model performance 5](#_Toc67411288)

[Chosen Regression equation 6](#_Toc67411289)

[Neural Network 6](#_Toc67411290)

[Development of models 6](#_Toc67411291)

[Model performance 6](#_Toc67411292)

[Neural network architecture of best model 7](#_Toc67411293)

[Decision Tree 7](#_Toc67411294)

[Development of models 7](#_Toc67411295)

[Performance of models 7](#_Toc67411296)

[Critical path 8](#_Toc67411297)

[Target path of interest 8](#_Toc67411298)

[Overfitting and limitations 8](#_Toc67411299)

[Research node 8](#_Toc67411300)

[Theory 8](#_Toc67411301)

[Settings 8](#_Toc67411302)

[Analysis 8](#_Toc67411303)

[Performance 9](#_Toc67411304)

[Analysis of the best model 9](#_Toc67411305)

[Conclusion 10](#_Toc67411306)

[Recommendation 10](#_Toc67411307)

[References and Bibliography 10](#_Toc67411308)

[Other Resources 10](#_Toc67411309)

[Appendix 12](#_Toc67411310)

[My Reflections on the process – What did I learn from this exercise? 12](#_Toc67411311)

# Summary

## Business Problem

## Data Mining Representation

## Methodology - data mining approach

## Data Exploration

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

Table 1 Summarising observations, measurement levels and data mining roles.

## Data partition creation of model sets

# Data Modelling

## Regression

### Modification Node

### Selection Methods

### Model performance

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Non Cu. % Res. Scope** | **Cu. Lift @ 10% depth** | **True -** | **False -** | **True +** | **False +** |
| Reg2 |  |  |  |  |  |  |
| Reg3 |  |  |  |  |  |  |
| Reg4 |  |  |  |  |  |  |
| Reg |  |  |  |  |  |  |

### Chosen Regression equation

Logit P =

## Neural Network

### Development of models

### Model performance

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Non Cu. % Res Scope** | **Cu. Lift**  **@10% depth** | **True -** | **False -** | **True +** | **False +** |
| Neural4 |  |  |  |  |  |  |
| Neural2 |  |  |  |  |  |  |
| Neural3 |  |  |  |  |  |  |
| Neural |  |  |  |  |  |  |

**Table 5 Neural network performance**

### Neural network architecture of best model

.

Figure 1 example neural network architecture

## Decision Tree

### Development of models

### Performance of models

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Non Cu. % Res. Scope** | **Cu. Lift** | **True -** | **False -** | **True +** | **False +** |
| Tree |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

Table 7 DT performance

### Critical path

### Target path of interest

### Overfitting and limitations

## Research node

### Theory

### Settings

### Analysis

### Performance

## Analysis of the best model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Non Cu. % Res Scope** | **Cu. Lift**  **@10% depth** | **Scope %** | **True –**  **%** | **False –**  **%** | **True +**  **%** | **False +**  **%** |
| Decision Tree |  |  |  |  |  |  |  |
| Neural2 |  |  |  |  |  |  |  |
| Reg14 |  |  |  |  |  |  |  |

**Table 8 summary results of the best performing models**

Figure 2 Cumulative lift chart of the best performing models

Figure 3 Non-cumulative lift chart of the best performing models

# Conclusion

# Recommendation

# References and Bibliography

# 

## Other Resources

## Appendix

Data Mining Roles

Workflow diagram

## My Reflections on the process – What did I learn from this exercise?