

Machine Learning

MOD006562

Faculty: Science and Engineering

School: Computing and Information Science

Academic Year: 2023/24

Trimester: 2

Contents

[1 Introduction 3](#_Toc125211706)

[2 Data Pre-processing and Preparation (15%) 3](#_Toc125211707)

[3 Statistical Analysis (10%) 3](#_Toc125211708)

[4 Preparing Data (05%) 4](#_Toc125211709)

[Train-Test Split 4](#_Toc125211710)

[5 Vectorization of Data (10%) 4](#_Toc125211711)

[6 Training and Evaluating Models (30%) 4](#_Toc125211712)

[5.1 [Model 1 – provide name of the model] 5](#_Toc125211713)

[5.1.1 Describe your model. 5](#_Toc125211714)

[5.1.2 Explain your evaluation design. 5](#_Toc125211715)

[5.1.3 Present and explain results. 5](#_Toc125211716)

[5.2 [Model 2 – provide name of the model] 5](#_Toc125211717)

[5.2.1 Describe your model. 5](#_Toc125211718)

[5.2.2 Explain your evaluation design. 5](#_Toc125211719)

[5.2.3 Present and explain results. 5](#_Toc125211720)

[5.3 [Model 3 – provide name of the model] 5](#_Toc125211721)

[5.3.1 Describe your model. 5](#_Toc125211722)

[5.3.2 Explain your evaluation design. 5](#_Toc125211723)

[5.3.3 Present and explain results. 5](#_Toc125211724)

[6 Results/Evaluation Analysis (20%) 5](#_Toc125211725)

[7. References 5](#_Toc125211726)

[DELETE this text:

**Report and Code:**

*To produce quality work in general, (i) give attention to report structure to make all sections readable and coherent in flow, (ii) use relevant academic references to support your content and (iii) use good coding practices.*

*Remove everything from the template that is in red colour or in [ ]*

# Introduction

[DELETE this text: A few lines explaining the task and structure.]

# Data Pre-processing and Preparation (15%)

[DELETE this text: Describe your dataset. An analysis of dataset and details of any data pr-processing and data wrangling needed to prepare the data for machine learning task.

# Statistical Analysis (10%)

[DELETE this text: In this section, you may describe characteristics of your selected data e.g., which features you selected from the complete dataset and which features were muted. It may be a separate section or just part of Section 2 where you may explain

* Initial exploratory data analysis (before pre-processing)
* Explain the pre-processing
* Further analysis of the data and its features
* Handling of missing data etc.

Hint: You may use tabular data and plots/graphs to visualise your data.]

# 4 Preparing Data (05%)

## Train-Test Split

[DELETE this text: Explain what your train-test split was. Based upon the data, you may explain some or all the followings:

* If you used Cross-validation technique
* What percent of training or test data you decided to use?
* Characteristics of your training data
* Characteristics of your test data
* Successful loading of data

Hint: You may use tabular data and plots/graphs to visualise your data.

]

# Vectorization of Data (10%)

[ Develop at least 3 different natural language vectorization (feature extraction) techniques. ]

# Training and Evaluating Models (30%)

[DELETE this text: In this section, you should describe which model/s you have chosen for your learning task. Give a brief introduction to the model and provide justifications for your choice.

]

## 5.1 [Model 1 – provide name of the model]

### 5.1.1 Describe your model.

### 5.1.2 Explain your evaluation design.

### 5.1.3 Present and explain results.

## 5.2 [Model 2 – provide name of the model]

### 5.2.1 Describe your model.

### 5.2.2 Explain your evaluation design.

### 5.2.3 Present and explain results.

## 5.3 [Model 3 – provide name of the model]

### 5.3.1 Describe your model.

### 5.3.2 Explain your evaluation design.

### 5.3.3 Present and explain results.

# 6 Results/Evaluation Analysis (20%)

Reflections and analysis of the results.

# 7. References

Fatima, A. 2021. Template for Machine Learning Assignment.