

**Postgraduate Certificate in Software Design with Artificial Intelligence**

**Data Mining and Machine Learning**

**Assignment 1**

Student ID: A00267948

Student Name: Daniel Foth

*Brief Description: This paper evaluates different data sets using regression, decision trees and kNN algorithms. Models are created and predictions are made using the test data from the data sets.*

Git: <https://github.com/DanielsHappyWorks/DM-ML-Module-Assignment>

Contents

[0. Introduction 3](#_Toc25683083)

[1. Regression 3](#_Toc25683084)

[1.1. Overview of the Problem 3](#_Toc25683085)

[1.2. Data Exploration (tables and graphs) 3](#_Toc25683086)

[1.3. Definition of Training and Testing Set 3](#_Toc25683087)

[1.4. Model Generation and Information 3](#_Toc25683088)

[1.5. Predictions for the test data 3](#_Toc25683089)

[1.6. Evaluation of the model(s) and conclusion. 3](#_Toc25683090)

[2. Decision Trees 3](#_Toc25683091)

[2.1. Overview of the Problem 3](#_Toc25683092)

[2.2. Data Exploration (tables and graphs) 3](#_Toc25683093)

[2.3. Definition of Training and Testing Set 3](#_Toc25683094)

[2.4. Model Generation and Information 3](#_Toc25683095)

[2.5. Predictions for the test data 3](#_Toc25683096)

[2.6. Evaluation of the model(s) and conclusion. 3](#_Toc25683097)

[3. kNN 3](#_Toc25683098)

[3.1. Overview of the Problem 3](#_Toc25683099)

[3.2. Data Exploration (tables and graphs) 3](#_Toc25683100)

[3.3. Definition of Training and Testing Set 3](#_Toc25683101)

[3.4. Model Generation and Information 3](#_Toc25683102)

[3.5. Predictions for the test data 3](#_Toc25683103)

[3.6. Evaluation of the model(s) and conclusion. 3](#_Toc25683104)

[4. References 3](#_Toc25683105)

# Introduction

# Regression

## Overview of the Problem

## Data Exploration (tables and graphs)

## Definition of Training and Testing Set

## Model Generation and Information

## Predictions for the test data

## Evaluation of the model(s) and conclusion.

# Decision Trees

## Overview of the Problem

## Data Exploration (tables and graphs)

## Definition of Training and Testing Set

## Model Generation and Information

## Predictions for the test data

## Evaluation of the model(s) and conclusion.

# kNN

## Overview of the Problem

## Data Exploration (tables and graphs)

## Definition of Training and Testing Set

## Model Generation and Information

## Predictions for the test data

## Evaluation of the model(s) and conclusion.

# References