**Digital Twin**

**Project**

**Data Science & Artificial Intelligence: DSAI**

**CU75074V2 – 2024/2025**

Date: <*DATE*>

Version number: 1.0

**Mistpedal Cuddlebeasts**

|  |  |
| --- | --- |
| **Name** | **Student number** |
| Jens Joosse | 00092511 |
| Hugo Berlanga |  |
|  |  |
|  |  |
|  |  |
|  |  |

**Document versions**

|  |  |  |
| --- | --- | --- |
| **Version number** | **Date** | **Modifications/status** |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Table of contents

[1. Introduction 4](#_Toc195876463)

[1.1. Describing data mining plans based on basic machine learning model (LO 1.1, 1.2) 4](#_Toc195876464)

[1.2. You add extra self-organized and/or external data sources to the data science process (LO 1.3) 4](#_Toc195876465)

[2. Approach 5](#_Toc195876466)

[2.1. Data description and validation (LO 2.1) 5](#_Toc195876467)

[2.2. Data merging & cleaning (LO 2.2, 2.3) 5](#_Toc195876468)

[2.3. Data preparation (LO 2.4, 2.5) 5](#_Toc195876469)

[Modeling and data analysis 6](#_Toc195876470)

[3.1 Model and metric description and test design (LO 3.1, 3.2) 6](#_Toc195876471)

[3.2 Results (LO 3.3) 6](#_Toc195876472)

[Conclusion 7](#_Toc195876473)

[4.1 Evaluate results in relation to the business objective and data science process (LO 4.1) 7](#_Toc195876474)

[4.2 Limitations (LO 4.2) 7](#_Toc195876475)

[4.3 Discussion & follow-up (4.3, 4.4) 7](#_Toc195876476)

[**Contribution and check-ins (indicator 7.3P)** 8](#_Toc195876477)

[**References** 9](#_Toc195876478)

[**Appendix 1 - <APPENDIX DOCUMENT TITLE>** 10](#_Toc195876479)

# Introduction

*<Write down a short introduction/background about the client and the problem (business understanding). Describe the situation, business objectives and business success criteria. >*

The HZ University of Applied Sciences in collaboration with the Lectoraat Data Science Domain Technology Water and Energy, assigned a project which is focused on improving the operation of the Waste Water Treatment Plant by creating a Digital Twin. This is a data-driven model that simulates the process of the Waste Water Treatment Plant, and mainly focusses on the “Phase 2” which is the chemical process in the treatment plant.

### Describing data mining plans based on basic machine learning model (LO 1.1, 1.2)

<*Describe your data mining plans based on what model you will be using and why (you can elaborate on details of the data in the approach section). Describe the relevant activities and why, and define your data mining success criteria. Remember to be as specific as possible in phrasing your DMGs and success criteria, and* ***why*** *you choose these.>*

### You add extra self-organized and/or external data sources to the data science process (LO 1.3)

<*If you are integrating other data, describe why you will use this data and where you obtained it. >*

## Approach

### 2.1. Data description and validation (LO 2.1)

<*Explain assumptions about your data, provide statistical summaries appropriate for the data type, use* ***and explain/interpret*** *appropriate visualization where applicable. Note that the information has to adhere to the formatting as stated in the Writing Guidelines. >*

### 2.2. Data merging & cleaning (LO 2.2, 2.3)

<*Describe and explain how your data was merged into a usable format, and explain how the data has become more logical and consistent. Describe and explain how and why imputation and scaling was applied, and what the consequences are on interpretation. >*

### 2.3. Data preparation (LO 2.4, 2.5)

<*Describe and explain the chosen dependent and independent variables, and describe and provide evidence for newly created variable (either by one-hot encoding or discretizing, and the labeling of said variable(s). Describe and explain how this contributes making your data available or more suitable for your chosen model technique.>*

## Modeling and data analysis

### 3.1 Model and metric description and test design (LO 3.1, 3.2)

<*Explain the test design, including argumentation how this prevents overfitting. Explain the modeling technique used in detail, how it works and how the output can be interpreted. Include the appropriate metric.>*

### 3.2 Results (LO 3.3)

<*Show, explain and* ***interpret*** *the metrics, and assess the model results (metrics) to your defined success criteria. Explain what the outcome means. >*

## Conclusion

### 4.1 Evaluate results in relation to the business objective and data science process (LO 4.1)

<*Evaluate and discuss success criteria with business objective and data mining goals, including an explanation why the goals have been met or not. Explain in laymen terms what your results mean and what they imply in relation to the business objectives. >*

### 4.2 Limitations (LO 4.2, LO 6.4D)

<*Describe and explain limitations of your project. Review the process and which lessons were learned.* ***Stick to the lessons learned that are relevant for your project, not things you personally thought were difficult.*** *>*

### 4.3 Discussion & follow-up (LO, 4.2, LO 6.4D)

<*Describe your advice and provide well-substantiated follow-up actions, derived from your results and conclusion. Be specific – for example, “we need more data” is too general/superficial. >*

# **Contribution and check-ins (indicator 7.3P)**

*<Record the attendance of teammates of the Wednesday check-ins and write down (briefly) their respective contribution for that week. Fill in the names of each group member in a column. Put an X if the student was not present that day, or if they did not contribute. >*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Week | Student1 | Student2 | Student3 | Student4 | Student5 |
| 1  (23-4-2025) |  |  |  |  |  |
| 2  (7-5-2025) |  |  |  |  |  |
| 3  (14-5-2025) |  |  |  |  |  |
| 4  (21-5-2025) |  |  |  |  |  |
| 5  (28-5-2025) |  |  |  |  |  |
| 6  (4-6-2025) |  |  |  |  |  |
| 7  (11-6-2025) |  |  |  |  |  |

# **References**

*<If you have used external literature during your work, refer to the source in the text with the note the source below in the bibliography in alphabetical order in APA documentation style.* ***Use the Word plug-in for this, and you can use Zotero as a bibliography manager during your project (see writing advice document).*** *For help with APA documentation, see the following examples for notation of books, journal articles and websites with 1 author. For additional examples see: https://www.scribbr.com/category/apa-style/ >*

Author Surname, Initials. (Publication year). *Title of the Book* (Edition ed.). Publisher.

Author Surname, Initials. (Publication year). Title of the article. *Title of the journal, Volume*(Edition), Pages. <https://doi.org/DOI>

Author Surname, Initials. (Publication year). *Title of the online source*. Website. https://URL

# **Appendix 1 - <APPENDIX DOCUMENT TITLE>**

*<In the appendix you can add additional evidence that you refer to in a previous chapter. You are allowed to include multiple appendices of course, with a maximum of 10 pages for all appendices combined.>*