­­­­­Personal Development Report

Thomas van der Molen

S7-AI-A-RB

|  |  |
| --- | --- |
| **Project Information** | |
| Project members | Thomas van der Molen (4168003) |
| Project name | Personal Development Report |

Table of Contents

[Introduction 3](#_Toc151587766)

[1. Data Preparation 4](#_Toc151587767)

[2. Data analysis & model engineering 6](#_Toc151587768)

[3. Targeted interactions 8](#_Toc151587769)

[4. Future orientation 10](#_Toc151587770)

[5. Investigative problem solving 12](#_Toc151587771)

[6. Personal leadership 14](#_Toc151587772)

[7. Graduation Preparation 16](#_Toc151587773)

[Retrospective 18](#_Toc151587774)

[Conclusion 18](#_Toc151587775)

# Introduction

My name is Thomas van der Molen, I am 21 years old and have previously done Mavo and Havo before coming to Fontys to study HBO ICT & Software Engineering.

I am currently working for the software company Author-e where I lead several different software projects.

I have done Advanced Software in my previous semester, and while I still want to focus on Software I think there is a lot to learn within the domain of AI, hence I decided to do Advanced AI as well.

During this semester, I want to learn more about the underlying algorithms power large AI models.

All previous semester I have been able to achieve an Outstanding grading, so I will try to achieve Outstanding this semester as well. However, I did recently learn that getting Outstanding every semester does not actually directly affect graduating with Cum Laude and for this reason I have put less importance on it.

# Data Preparation

You can **enrich/extend** and **clean** datasets, and are able to investigate and argument the use of **data storing methods**.

Evaluation 1: I think this learning outcome is currently at an **Orienting** level.

Evaluation 2: I think this learning outcome is currently at a **Proficient** level.

**Evidence Evaluation 1**

During the start of the group project, I have spent a lot of time researching varying DevOps tools specifically for machine learning projects, especially DVC for which I have created a research document, with guide for fellow students and have implemented my findings for data / model / experiment -management in the group project.

Furthermore, I have started performing data analysis on the data of the group project together with my project group, and have done initial dataset exploration for smaller datasets used during exercises which I have extensively discussed with Technical tutor (Olaf).

**Evidence Evaluation 2**

Since the last evaluation, I have created together with the rest of the project group the final data preparation notebook, this notebook contains the data merging techniques I had utilized previously in my EDA for One-class SVM and shared with the rest of the group. In this merging process, the datasets are grouped on specific identifiers, effectively squashing the data into less records, this process also creates new features such as average/median, deviations, timespans etc. that could enrich data records with timeseries-like features that from my research should be good indicators for predicting anomalies.

In the pre-processing file, I also create many versions of the datasets, including final Train/Test split datasets that can be used to standardize the data used when training/evaluating so that the various models can be compared more closely with each other. All the saved models are stored in DVC that I had previously set-up, allowing group-members and cloud pipelines to easily fetch the datasets needed (which with these enrich and computed datasets are starting to reach the 1GB file size which is already larger than GitHub’s 100MiB file push limit).

Feedback from Olaf confirms that the processing done was very well (note this feedback is of the whole data-preprocessing notebook which includes work of other students Fig 1), but more information of the reasoning behind the processing steps (EDA) could still be added.

A screenshot of a phone

Description automatically generated

Figure 1 Olaf Feedback SUE data-preparation

Furthermore, I have also started looking into how to convert web pages (HTML/CSS/JS) into reinforcement learning environments, for my open-program/data challenge. This work was shortly covered in my data challenge plan, on which Qin gave positive feedback on (Fig 2).

A screenshot of a phone

Description automatically generated

Figure 2 Qin data challenge plan feedback

Due to the vast amount of work I have done on the data preparation for group project and my own projects, I think I have shown a sufficient level of knowledge and execution within this field.

# Data analysis & model engineering

*You can* ***apply*** *different types of machine learning models and tune the* ***hyper-parameters*** *of these models. Furthermore, you can* ***evaluate*** *the results from a trained model and can perform* ***other data analysis techniques****.*

Evaluation 1: I think this learning outcome is currently at an **Orienting** level.

Evaluation 2: I feel that this learning outcome has improved to a **Beginning** level.

**Evidence Evaluation 1**

Besides the workshops offered during this semester, I have explored models of which I had limited understanding such as CNN, for which I had actually found very unique issues while evaluating the model, which I have extensively discussed and further explored based on feedback from the Technical tutor (Olaf).

**Evidence Evaluation 2**

To continue from Evaluation 1, I have spent even more time exploring issues I had with my ANN and CNN model exercises, however this time I used a more methodological approach with hypothesizing possible issues and utilizing XAI to prove these hypotheses, the effectiveness of this approach can be seen by the feedback left by Nico on my XAI exercise (Fig 1), with my ideas being earlier confirmed by Nico in my CNN exercise itself (Fig 2).

A screenshot of a test

Description automatically generatedA screenshot of a chat

Description automatically generated

Figure 2 Feedback (Highlighted) Nico CNN - Exercise

Figure 1 Feedback Nico XAI - Exercise

For my group project, I have also explored several kinds of outlier detection algorithms, mainly focusing on One-class SVM (including an entire research document) which I have discussed in great detail with my technical tutor (Olaf) and transformers for which I have done a significant amount of preliminary research and asking a lot of questions to Olaf about the topic after his NLP lectures, which from the feedback received during the meetings seems to be going in a good direction.

I also spent a lot of time working on the RL exercise, as I will need the knowledge and experience within this context for my open/data challenge project. During the RL exercise, I spent a significant amount of time exploring several environments and different types of learning models (notably Q-tables and deep learning models), investigating how certain techniques can be implemented and combined and why they work in the first place. All my efforts did result in some interesting/good reinforcement models that I fully understood and could explain, which Iman seemed to also be positive about according to his feedback (Fig 3).

A screenshot of a phone

Description automatically generated

Figure 3 Feedback (Highlighted) Iman RL – Exercise

From the evidence shown so far I think that I have proven a beginning and closing in on proficient level of knowledge within the data analysis and model engineering scope.

# Targeted interactions

*You can* ***clearly communicate*** *with readers/audience, keeping into* ***consideration*** *the knowledge of set audience and the medium used to communicate information.*

Evaluation 1: I think this learning outcome is currently at an **Orienting** level.

Evaluation 2: I think that I have proven a **Beginning** or even **Proficient** level for this learning outcome.

**Evidence Evaluation 1**

During the early parts of the project, I have created a Business Proposal in collaboration with my project group, for which feedback from the Process tutor (Martijn) was very positive, keeping the audience in mind. Furthermore, I have created a DVC research intended to document my findings and be able to share this with my fellow students for them to use it as a guide, also providing a README in my group project with instructions of how to use DVC within our project.

During the initial meeting with SUE, I have also taken a leading position in communication both during the meeting and general communication via email/google chat in which I showed my ability to communicate clearly with stakeholder.

**Evidence Evaluation 2**

I am still taking a leading position in my group project meetings, making sure that I am well prepared and know what general topics will have to be discussed. This became especially useful when having a meeting with a researcher from TU/Delft about the paper he wrote, as I was the only one to have actually read it before hand, this allowed us to ask insightful questions and make communication significantly easier due to the fundamental understanding on the (quite advanced) topics discussed in the research paper.

I also generally take the lead in meetings with the group project stakeholder, which the stakeholder does not seem to have any issues with, as I try to ensure that everyone is on the same page and that the stakeholder knows what we are talking about (which up until now seems to have worked, as I have only gotten positive feedback from the stakeholder and never got any requests to try and further explain a topic), this was also put into the group feedpulse by Martijn (Fig 1)

**A white background with black text

Description automatically generated**

Figure 1 Martijn Group feedpulse week 8

Besides my group project, I have also created a lot more documentation (including documentation on DVC and using GPU with TensorFlow on windows which seemed to have been useful resources for fellow students). Based on feedback received from teachers on my exercises I have also tried to keep improving my documentation within notebooks, which for the most part seemed to have been received positive, as shown by the feedback received from Olaf on group project work (on which I spent a lot of time getting a good structure while including work from all group members) (Fig 2), and my data challenge plan, which I was able to effectively go through and discuss with Qin (Fig 3).

**A screenshot of a phone

Description automatically generated**

Figure 2 Olaf feedback Exploratory Data Analysis

A screenshot of a phone

Description automatically generated

Figure 3 Qin feedback Data challenge plan

From all my work done so far to improve upon feedback received for Targeted Interactions, I think that I have proven my proficiency within this learning outcome.

# Future orientation

*You can* ***identify and explore*** *the* ***context*** *in which your project takes place, and you can approach this from* ***multiple perspectives*** *within the machine learning domain.*

Evaluation 1: I think this learning outcome is currently at an **Orienting** level.

Evaluation 2: I think this learning outcome is currently at a **Beginning** level.

**Evidence Evaluation 1**

For the set-up of the group project, I ensured that a Business proposal was created that could clearly show our understanding and vision of the project, to ensure that no unexpected issues will be encountered later on, including setting up effective communication channels with the stakeholders.

Secondly I have spent a lot of time researching and setting up a robust DevOps environment for the group project that the stakeholder company can adopt even after the project in finished, using best practices and tools explored in my DVC research.

**Evidence Evaluation 2**

Since the previous evaluation, I have spent more time exploring both my group project and data challenge. For my group project, I have reached out to a researcher at the TU/Delft who co-wrote the research paper a lot of our work is based on. To ensure that a meeting with the researcher would be worthwhile for everyone involved, I had read both the paper in question and (most of) the related papers, which he cited in his work or had written as well to extend the research done. This proved extremely useful as it has given me far greater insight into the topics discussed and allowed me to ask far better and targeted questions to help the group project further.

For the group project, I had also done more exploration into the data as I had discussed possible ways of enriching the data for specific models (such as adding time-series data to a model that normally does not account for time dependent data), which ended up in the group project’s EDA and data processing, which both technical tutors mentioned showed potential (Fig 1).

A screenshot of a phone

Description automatically generated

Figure 1 Feedback (Highlighted) Olaf Exploratory Data Analysis

Furthermore, I have also spent time during my open-program and start of data challenge to explore the required context of my RL-powered integration test project, this has culminated in a project plan for the data challenge, which after discussing with the data challenge tutors was received very positively (Fig 2).

A text on a screen

Description automatically generated

Figure 2 Feedback Qin data challenge plan

The evidence shown so far is a good beginning for proving proficiency for this learning outcome, possibly showing a proficient level with more evidence.

# Investigative problem solving

*You can Identify a problem and create a* ***hypothesis*** *and* ***research question*** *and can do this using a sound* ***research methodology****.*

Evaluation 1: I think this learning outcome is currently at an **Orienting** level.

Evaluation 2: I think this learning outcome is at a **Proficient** level.

**Evidence Evaluation 1**

I have created a research document, exploring DVC. This research document is done using the DOT framework methodology, including the machine learning extension specifically added for the machine learning context.

This research document has been shown to the Process tutor (Martijn) who was positive about the methodology and research done.

**Evidence Evaluation 2**

After the first Evaluation, Martijn gave me a Beginning instead of my self-graded Orienting, due to the amount of work I had already done within the group and core program (Fig 1)

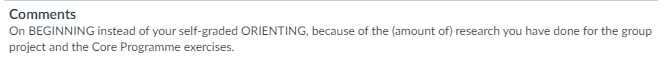


Figure 1 Martijn grade reasoning for Investigative Problem solving

Since that Evaluation, I have done a lot more research into many different topics such as the Reinforcement learning exercises, which also extend into my open/data challenge. Iman ended up giving me very positive feedback on my approach to the Reinforcement Learning exercise, specifically on my methodology for problem-solving and approaching a problem (Fig 2).

A screenshot of a computer

Description automatically generated

Figure 2 Iman feedback RL - Exercise

I used the same approach for all my other exercises, where if I ran into an issue, I would research the problem and create a hypothesis of what is causing it, with sometimes even being able to solve the issue afterwards. Two notable instances, were my CNN exercise and XAI (as I was able to prove some hypothesis of my other models during my XAI), my successful application of my approach can be seen in the positive feedback received from Olaf (during meetings in person) and Nico (See Fig 3/4 and in person discussions).

A screenshot of a phone

Description automatically generatedA screenshot of a text message

Description automatically generated

Figure 4 Nico feedback XAI – Exercise

Figure 3 Nico feedback (highlighted) CNN - Exercise

Furthermore, I had also created a plan for my data challenge, with contained a condensed and simplified version of my research/results done so far on the topic as a part of my approach, for which I also received positive feedback for by Qin (Fig 5).

A screenshot of a phone

Description automatically generated

Figure 5 Qin feedback Plan: Data Driven Innovation Challenge

Due to all the work done since my last evaluation, and the proficiency shown throughout many of my projects, I feel confident in my Proficient level for investigative problem solving.

# Personal leadership

*You are* ***proactively*** *looking to improve your own* ***development*** *and* ***learning****, while taking a* ***leading position*** *in your projects.*

Evaluation 1: I think this learning outcome is currently at an **Orienting** level.

Evaluation 2: I think I have proven a **Proficient** level of proficiency for this learning outcome.

**Evidence Evaluation 1**

From the start of the semester, I have tried to ensure that my work can properly cover the learning outcomes set out for this semester. Secondly, I have taken a leading position within the group project, often leading meetings both within the group and with stakeholders.

Furthermore, I have spent significant time performing research and exploring machine learning within my own time to try and further understand topics that might not be covered by the normal curriculum (excluding the Open Program). I have also tried to be very proactive in asking for feedback on my work done, giving me valuable knowledge and ideas from both tutors and fellow students.

**Evidence Evaluation 2**

I have tried to have meetings with tutors and other stakeholders as often as possible to get feedback on my progress so far, which has helped me quickly adapt and learn when for example making incorrect assumptions about model performance.

During the group project, I am still taking a leading role in most meetings and outside of the meetings, including making sure assignments are submitted on time, and feedback/standups are done regularly, allowing the group project to run smoothly and make constant progress.

During this semester, there were two things that I wanted to improve on further with the first one being ethical concerns. As I knew this was a topic that would be covered later during this semester (and was covered for the Software domain during advanced software), I decided to get a jumpstart and already investigate potential ethical concerns in my open program, which was then transferred to my data challenge plan. I discussed the ethical concerns with Qin and received positive feedback on my approach (Fig 1) and was offered tools to look at further.

A close up of a text

Description automatically generated

Figure 1 Qin Feedback data challenge plan

One thing that I have always been bad at, is approaching people/companies which during my preparation for future after Fontys, was identified as one of the biggest potential failing points if I were to start working as a ZZP’er. To try and improve on this skill, I decided to go to the Career day at Fontys TQ, for this I had created a QR-code that linked directly to my CV and had created (sort of a minigame) challenge of getting as many companies as possible to scan the QR-code and look at my CV. As expected, at the start of the day, I was very hesitant to show it as I could not naturally work it into the conversations. But after talking with InfoSupport who were extremely excited when I mentioned I had a CV that they could get via a QR-code (instead of a piece of paper), I became far more confident in not only just asking people to view my CV, but in talking with companies in general (eventually leading 10+ companies to have scanned my QR-code and viewed my CV, which has now caused most companies to contact me to mention this fact meaning that it might have actually helped me stand out).

Due to all the progress and continued leadership I have shown during this phase, I think I have shown a sufficient level of personal leadership.

# Graduation Preparation

*You create* ***opportunities*** *for finding a* ***graduation internship****, that aligns with your* ***personal needs*** *and the* ***school’s requirements****.*

Evaluation 1: I think this learning outcome is currently at a **Beginning** level.

Evaluation 2: I think I have achieved a **Proficient** level for this learning outcome.

**Evidence Evaluation 1**

Before the start of the semester, I have already spent a lot of time building up an understanding of what I want to do after graduation, including interviews and domain explorations and improving my outwards appearance for companies by improving my LinkedIn, GitHub, CV and creating a portfolio website.

I already have created a full planning for finding a graduation internship, including fallback plans which can be found in my Open Program Plan.

The work done so far has already seemed to pay of with me getting semi-daily emails and LinkedIn messages from new companies and previously approach companies both for working opportunities and internships.

**Evidence Evaluation 2**

Since the first evaluation, I have done a lot of work on finding a graduation internship (most of it directly following the previously made planning).

First thing I did was create a spreadsheet containing a list of companies I am interested in, either from LinkedIn or the ASAM portal with all required information. I have been using this spreadsheet to keep track of the status and remind me of my previous opinion of a company (Figure 1).

Internship candidates spreadsheet


Figure 4 Internship candidates spreadsheet

From this spreadsheet, I had reached out to my T1 (green) companies, sadly these companies did not have (or want me for) an internship, so I moved on to my T2 (blue), while also responding to all companies that had contacted me on LinkedIn (even the ones offering me a job instead of internship).

During this period, I was still trying to get an internship with SpendLab (my nr1 candidate), as they were not responding to my emails making me contact them via other means and eventually getting into contact with the previous recruitment manager who had apparently told their successor to reach out to me (which he never did). After getting into contact with the new recruitment manager, he told me they had filled all internship spots, but forwarded me to a different company (Waves Process Intelligence) (while he mentioned he did that because of my excellent portfolio, I have a feeling that he might also feel bad that I could have gotten the internship if he had just replied to me earlier).

After corresponding with Waves Process Intelligence, and having an in-person interview I was offered 3 possible projects to do for them (which all of them are very interesting, and I even offered to share them with more Business/AI students), meaning that I now have a company I can do my graduation internship at. The company however, is very small so for now I am still looking for other companies knowing that I have a fallback company that I can always do the internship at.

I also went to the Career-day at TQ, here I talked with many companies (including some that did not respond to my emails…). As preparation for the Career-day, I had made my CV into a QR-code and shared it with companies during the day (companies were surprisingly extremely excited about this), which got me an interview with InfoSupport BDO and Alten in the coming week, and correspondents with others.

After all this work, I have been able to find a fitting graduation internship that I would like to do, but have also been able to get into contact and be offered interviews with more companies (as I would still prefer to do an internship at a “large” company to get experience within that environment).

# Retrospective

# Conclusion