

Personal Development Report

Thomas van der Molen

S4-AI41

Project Information	
Project members	Thomas van der Molen (4168003)
Project name	Personal Development Report
Version	1.2

Table of Contents

Version History.....	2
Introduction	3
Data Preparation.....	4
Data analysis & model engineering	5
Reliability and transparency	6
Targeted interactions.....	7
Future orientation.....	8
Investigative problem solving	9
Personal leadership.....	10
Internship Preparation.....	11
Retrospective	12
Conclusion.....	13

Version History

Version	Date	Change
1.0	15-02-2022	Created document
1.1	15-03-2022	Added a personal Introduction
1.2	13-04-2022	Added Evidence gathered from feedback from iteration 0 and partially iteration 1

Introduction

My name is Thomas van der Molen, I am 19 years old and have previously done Mavo and Havo before coming to Fontys to study HBO ICT & Software Engineering. While I have always been interested in IT, I did not have a background in software before coming to Fontys. During high school I took an IT class, but this did not give me any extra knowledge in the field of software engineering that I had not learned before from myself.

I have chosen to do a specialization in AI because I am very interested in the subject and think it would round out my software background nicely.

I think there is a future in the AI sector and am very interested in where this area might go in the future. I also want to learn more on the process of creating a model and training it, because I think the algorithms used are very interesting.

Because I already have a background in software, I assume that the machine learning part will come easier to me and for this I also want to spend a lot of time on data preparation because I think that I could learn a lot in this part of AI.



Data Preparation

*You are able to **prepare** and **store** a given dataset in such a way that it can be used in your data analysis and/or modelling.*

Clarification:

Preparing a dataset consists of extending it with additional data and cleaning the data according to theories of data quality, in such a way that the process of cleaning and preparing those data is repeatable, transparent to others, and the results are suitable for data analysis and/or modelling.

Storing a dataset includes investigating cloud solutions and arguing whether they should be used.

Evidence

Thank you for your submission! I think you did a good job, very extensive data analysis. I would expect that for such big research you would have a bit longer conclusion. Also, I think what will help you is to write down some subquestions

Pencheva, Sabina S., 15 Mar at 10:15

From the quote above gotten from my Iteration 0 delivery on canvas by Sabina, she mentions that I did very extensive and well-done data analysis. She had no direct comments or changes that had to be made to this part of my iteration, only having improvement feedback on my research components for the analytic approach and reporting.

This feedback was very positive and showed that I am on the right track with my data preparation and EDA.

Data analysis & model engineering

You are able to **apply** machine learning algorithms and **other data analysis techniques** to a prepared dataset.

Clarification:

Applying consists of training of different types of models (classification, regression, etc.), and evaluating the results with respect to recall, precision, accuracy, cross-validation, etc. as well as tuning hyper-parameters.

Other data analysis techniques are for example: descriptive statistics, derived columns, forecasting, trend analysis, clustering, etc.

Evidence

Hi Thomas a few things to note: First off, I believe you went a little beyond the idea of the "iteration zero" and sort of included and iteration one" as well with the RANSAC. Which is ok, because there is still enough time for further iterations. I like that you reason on what the results are and why they are the way they are, so that is good! 👍 For me this challenge is a GO

Michielsen, Bas B.S.H.T., 15 Mar at 16:37

The above text is a snippet taken from feedback received on my Iteration 0 by Bas, he goes on to discuss possible food for thought which I dove deeper into during Iteration 1.

Bas mentions that I might have done a little too much for the scope of Iteration 0. However, Bas went on to explain that he thinks I did good research and had given valid reasoning for my results during the modelling phase.

I think this feedback was very positive combined with giving good future points of interest for further iterations and Bas also directly gave me a GO on this challenge, indicating that I have shown clear potential with my project.

Reliability and transparency

*You are able to address **reliability** and **transparency** aspects during data analysis and/or modelling.*

Clarification:

Reliable means that conclusions are supported by methodologically acquired and evaluated results, recommendations make sense in regards to the domain knowledge, and used sources are referenced appropriately.

Transparent refers to the process being clear to such a degree that it is reproducible, results being explainable to humans and based on decision making that is considered fair, whilst eliminating bias.

Targeted interactions

You use **appropriate communication** to address your audience considering your role, your audience and the medium to convey your message regarding the **approach** and **results** of your data analysis and/or modelling.

Clarification:

Appropriate communication means reporting and/or presenting the approach, process or outcomes of a data analysis in a methodologically sound way. You can reflect on the effect of your communication and based on that reflection you define steppingstones ahead on the task, on the role and on the projected results.

Interaction about data analysis is based on a systematic **approach** (such as the IBM Data Science Methodology) and comprises 4 phases/components: 1) project proposal 2) data preparation 3) data analysis/modelling 4) results and evaluation.

Based on your message, your position and the person(s) you address, you choose the right channel and format to communicate **results**, including appropriate data visualizations (such as data stories, infographics, or (a set of) static or interactive plot(s)).

Please explain further I do not have a clear Idea how this learning outcome can be separated from others.

Future orientation

*You are able to **assess the context** of data analysis and/or modelling from **multiple perspectives** in order to pursue this project in a sustainable manner.*

Clarification:

Assess the context means the organizational and societal environment in which the data analysis project takes place. You show that you can identify the hallmarks and roles of the environment of the assignment and have a keen eye for a future-oriented, sustainable embedding of your work in an organization and society.

Multiple perspectives include social and ethical considerations, law compliance, organizational data maturity, alignment with sustainable development goals, recognizing own boundaries and those of others and acting accordingly. Reflecting on ethics and governance of AI-based automated decision-making will be an important and integral part of your learning process.

Evidence

Hey Thomas, It seems you sufficiently dove into the domain of movies. Could be a little more extensive, and too bad you can't find an expert, but maybe you could then add some "literary source"(?) (<-maybe a piece about what makes the "world of movies" tick?) to use as an expert? If you can't really find that, don't put too much more time in that and go on with EDA-ing. Good start so far!

Bloks, Danny D., 8 Mar at 14:12

The feedback shown above was given by Danny on my Iteration 0, he starts this feedback by mentioning that I sufficiently dove into the domain of movies, Danny goes on further highlighting the problems that I have had trying to get in contact with a domain expert. It is also mentioned that I could add Literary sources to improve my understanding.

As we discussed recently, in this way the Proposal Phase seems sufficiently displayed in this document and your process. So good job. For the Provisioning Phase you should ask Sabina. Good luck!

Bloks, Danny D., 5 Apr at 14:36

During Iteration 1, I continued looking for domain experts with the help of my semester coach, however I also took Danny's advice and tried finding more different angles to approach my domain from and adding literary sources.

As said in the feedback above from Iteration 1, I seem to have done this to a satisfactory degree, giving me no real points of improvement and mentioning that I can move on to the next phase.

From this feedback I think I have done plenty of research into the context and effects of my project and showing a good investigation into the different perspectives of my domain and project.

Investigative problem solving

*You are able to critically consider a data analysis and/or modelling project from various perspectives, **identifying problems**, finding an **effective approach** and coming up with appropriate solutions.*

Clarification:

Identifying the problem means you can formulate a clear hypothesis and research question in order to determine the aim of solution using an inquiring mindset. **Effective approach** means you compose and pursue applied research methods (such as the DOT framework) and approaches based on reliable and verifiable sources. You are able to methodically and creatively find answers to applied research questions, considering alternatives and critically analysing your own and others' line of reasoning.

Evidence

Also, I think what will help you is to write down some subquestions

Pencheva, Sabina S., 15 Mar at 10:15

During Iteration 0, Sabina gave me the feedback to add sub questions besides my main question to give me a better Idea of what to expect and what could impact my challenge. I added these sub questions afterwards and did indeed do more research into different aspects of my domain and possible factors that could affect my results.

One of these factors were other features that I had not considered during my Iteration 0 initially such as actors impacting the performance of a movie, I had written this out with my reasoning for this. Bas had later given me feedback on this which is displayed below.

I do agree with your statement that perhaps enriching the dataset will give you more usable features.

Michielsen, Bas B.S.H.T., 15 Mar at 16:37

From this it seems that I have done a good job at being able to formulate and convey my reasoning and hypotheses made, with the example above given as one indication of this process.

Personal leadership

*You show an **entrepreneurial mindset** regarding the data analysis and/or modelling project and your personal development, while being aware of your own **learning capacity** and keeping in mind your **professional ambitions** in the field of Artificial Intelligence.*

Clarification:

Entrepreneurial mindset includes being aware, seeing opportunities and seizing them, motivating oneself and others, being able to profile oneself, a team and others.

Learning capacity means guiding your own development and study progress, showing leadership and taking responsibility, enhancing ones' own learning capacity, demanding and giving active feedback, all with respect to the learning outcomes.

Professional ambitions means you are examining what type of professional you want to be in the long term, which field and type of position you aspire to and how you can stand out from others in field of AI/data.

Could this be feedback I ask from my semestercoach? I feel like getting written feedback for this is pretty hard.

Internship Preparation

You create chances to acquire and define an internship assignment based on a match between your ambitions, the school's requirements and the field of expertise related to your profile or specialisation.

Retrospective

Conclusion