

PERSONAL DEVELOPMENT REPORT

By Mihail Kenarov

I am Mihail Kenarov, a semester 4 student at Fontys, who is interested in AI. In this document I will record the steps that I take throughout this semester, as well as the progress that I make in the sphere of AI. This will include the social impact that it has, my understanding of data analytics and investigative analysis and finally my knowledge of machine learning and how it works and the way that it can be used. This will be a living document which will be updated once time an evaluation phase occurs.

1. Learning outcome 1: Data Preparation and Analysis

Clarification:

Data preparation and analysis of data is the process of aggregating and preparing given dataset as well as other open datasets and use them to analyse and identify the opportunities for predictive analytics.

Aggregate means acquiring data from a variety of different sources and in different formats and putting it together into a meaningful larger total dataset.

Prepare consists of 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.

Data analysis implies amongst others: descriptive analytics, statistical overviews, derived columns, trend analysis, etc.

Opportunities for predictive analytics can be identified by finding correlations between features, principle component analysis, summarization, anomaly detection, etc. and include an impact forecast.

Second Evaluation: week 10

After 'dissecting' the first dataset that I chose for my Individual challenge, I began to see some troubles that occurred for the features that I chose for the project. For example, the 'ratings' column which at the time I considered of major importance, was missing many values. After having a discussion with our teacher – Hans Konings, we decided that it would be a good idea to check if there are other datasets that could be of use for the problem, that may solve this issue.

While doing so I found a new dataset that I am going to use. It is full of different columns, is more versatile and last, but not least, is more up to date.

After looking into it, the same problem occurred, but I this time I have other values to work with – "wishlist, playing, playing ...". After a bit of research and some I'd say logical thinking, I decided to drop the ratings column, because the missing rows could be of either it being an indie game that not a lot of people know of, it might be an older game, etc..

While doing this I also continued to take courses that can help me understand better what I need to do with data.

During this time we also received some feedback for the group project. Luckily, I am with a team who knows what they are doing and they are willing to show me what are some of the ways that they looked into the data, what they did with it, why they undertook those actions and also created visualizations for easier understanding.



Hans
Konings

15 April 2024 | [Semester Project - Iteration Zero \(w7\)](#) ↗

Hi TodayAI, Nice to bring up Theo's case and what you already did so far, and the next steps you're going to take! Visuals clear for target audience with only a small remark regarding the header and subsets. Presentation time is well balanced.

Suggestion: after presenting, it can be better to first ask the client whether they have questions before asking your own questions. As the client mentioned, please pay attention to one dataset at a time before merging everything together. 👍



Hans
Konings

15 April 2024 | [Data Visualizations Advanced](#) ↗

Hi TodayAI, great improvements were visible in the dashboards you made in the second attempt. You applied the knowledge of the lecture regarding the forecast scheme, the awareness of color mapping (and the lack of it) and strip the styles to focus more on the insights you want to provide. Really great work and a role model how to create and upgrade your own work. Well done! 😊



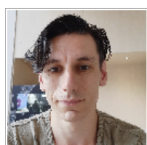
Self assessment: Beginning

I found my previous mistake, chose another dataset, looked through it and analysed it, while also making it more suitable for the model

First Evaluation: week 5

I have looked up different datasets for an individual challenge and have gone through their data and accessed if they are eligible for the creation of my first project in the sphere of AI. I have also checked the legitimacy of their data, the source they have gotten it from and if what I found in it is similar to the data that other sources offer. While doing so I have also exercised what I have learned and tried to implement it in my individual challenge.

While doing the exercises of week one, I received feedback from our teacher – Hans Konings



Hans
Konings

27 February 2024 | [Intro to Data Provisioning](#) ↗

Hi Mihail, correct observation that the notebook is lacking anything related to model engineering and explainable AI. For Data Preparation and Analysis, I do agree that it is quite basic, even though there some descriptive analytics to be found like min-max values per feature etc. Would you have more ideas how to improve it further? (food for thought for the moment you work on your challenge 😊)

Self assessment: Orienting

Because I have looked through different datasets, checked their legitimacy and I am still practising on how to analyse and work on the data.

2. Learning Outcome 2: Model Engineering

Clarification:

Model engineering is the ability to use findings from data to preprocess data, apply machine learning algorithms and evaluate the quality and usefulness of produced models, for a defined domain.

Findings from data analysis implies that your choice of data sources and feature selection is based on opportunities for predictive analytics that you previously identified.

Preprocess refers to applying systematic ways like feature selection, encoding, scaling, etc. of turning raw datasets into formats that are more suitable for model training.

Apply consists of training of different types of models like classification, regression, etc., as well as tuning hyper-parameters.

Evaluate means judging the results of machine learning with respect to recall, precision, accuracy, cross-validation, over/underfitted etc.

A defined domain refers to the fact that your evaluation must address the problem and impact definition as given by the domain stakeholders, and evaluation metrics must be translated to be meaningful to them.

Second Evaluation: week 10

Since the reset of my dataset, I have also looked more into the different kinds of recommendation system and the algorithms used in them. I am considering adding FG*IDF algorithm, so that I can use it on the description to possibly find similar games. By doing this we can see the words and how many times they are in a current sentence/description of a game. Afterwards, we are vectorizing the results

After doing so, we can use the sigmoid kernel that gives us the opportunity to compare the vectors of sentences with each other. This will show us the results in a scope between 0 and 1. The closer the number is – the more similar a game is to another.

I also had a talk with our teacher - Priyanka Darbari, after the swapping of the dataset and we discussed that I am not behind, however I should discuss with her the new target variable and also the features selected for the new model.

I showed my Machine learning teacher what I did and here is the what we discussed:

“The way you explain your notebook is good. You apply the sigmoid kernel which is nice and you need to add graphs of the results and make the conclusion. In next iteration I would like to see the comparison between different machine learning model, including new feature selection.(With new data preprocessing). Add references about the research (Where I found the sources and why I am using it). Add a small introduction about iteration 0 in next iteration. What I did, what would be the result, what I am going to do in the future.”

Self assessment: Beginning

I have learned about other models, their uses, are there more suitable models for my individual challenge and how to implement them.

First Evaluation: week 5

The assignments that have been provided to us have been all made except the last one. I have a basic understanding of what they are, why are they used, what are they used for and also how can we apply them. Based on this I have chosen a the Nearest Neighbours Method (kNN), since I have decided that for my project, which is a recommendation system, based on the selected features it is a matter of classification. The idea is to show 3 games similar to the one that a user has chosen and see if the user will like them.

It does seem that I am going in the right direction as a whole for the project, however some things might be missing and I need to look deeper into the specified topics, mentioned from one of my teachers – Priyanka Darbari



Priyanka
Darbari

18 March 2024 | Challenge 1 - Iteration Zero [🔗](#)

👍 You explain very well the domain understanding with diagram but analytic approach needs more description and add your research question as well with explanation.

Self assessment: Orienting

Because I have looked at different projects, different datasets and decided that this particular model is useful for the current challenge, based on the features of the dataset.

3. Learning outcome 3: Explainable AI

Clarification:

Explainable AI is the ability to deliver projects that follow the three 'Explainable AI' principles of transparency, interpretability, and explainability.

Transparency means that the process by which the used input data results in prediction models is reproducible, reliably described and its decisions are motivated.

Interpretability addresses the possibility for humans to comprehend the project cohesion and results by making them comparable to the domain knowledge and baselines.

Explainability refers to the application of tools and methods that turn black-box models into grey/white-box models by having the model draw out its decision making process and/or describe its feature importance.

Second Evaluation: week 10

After thorough research on the topic, I have a better understanding of what are some of the pros and cons of recommendation system. What they lack in functionality, as well as the lack of ethical thought behind the whole problem with the business. How the businesses implement them, with “care” for the customers, while not caring what happens to the actual person, but just that the sales

must be happening. Not only are they just interested in us buying the products, the whole idea of the user, is that they are more so looked at as a “consumer”, not that much as a person.

During these last couple of weeks I have also been criticized about the way I give credit to the people and sites that I have gained my knowledge on the topics from. Because of this I am starting to understand, bit by bit, how to give better credit, to the places that have helped me gain more insight on the topics that I am interested in.

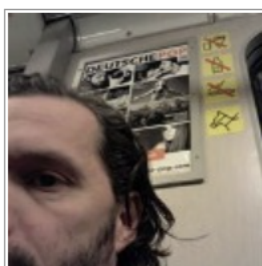
Self assessment: Beginning

I now know more about the pros and cons of the recommendation systems, where are they used, how are they used. The ethical problems with them, not only the functional ones, as well as how to give credit better, so that the places and people who have helped me get the recognition that is deserved.

First Evaluation: week 5

The exercises that have been given to us throughout the semester until now have been done either in group or alone. I have also done research on recommendation systems, how they work and what types there are. While doing so I also understood the importance of them and how they operate in the world. Without them many stores would not be able to sell as much product, no matter the type of product and the user would not be able to find products that could be of interest to them.

After doing the exercise about domain understanding and cognitive maps, me and my colleague, Aleksander Konstantinov, managed to receive this feedback as well, from our teacher – John van Litsenburg:

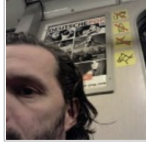


John
van Litsenburg

10 March 2024 | [2 Domain Understanding - General](#)

Nice work!

However, after the submission of my Project proposal, it does seem that I am missing some of the requirements that I need to do, to fully understand how to write such a document and it does show some lack of understanding from my part on the subject



John
van Litsenburg

21 March 2024 | Challenge 1 - Iteration Zero [🔗](#)

Hello Mihail, i think this proposal is only still half way there. I would like to discuss this further. You can continue with the topic but you now have to get more content into the proposal. So do not only talk about what you plan to do although this might sound contradictory to what a project plan is. But back this up with exploratory / provisional results; so the proposal and the work on iteration zero and further overlaps. For instance the proposal should already contain and display more concrete domain understanding and for instance research questions also pertaining to societal impact matters in this case of recommender systems. I prefer to discuss this face to face.

Self assessment: Unoriented

The reasoning behind this claim is the knowledge that I have on recommendation systems, as well as the exercises that have been done to solidify my knowledge on the subject, but despite that it does seem that I am missing some form of understanding on the subject.

4. Learning outcome 4: Professional standard

Clarification:

You show that you conduct work in accordance with an industry supported methodological approach (AI Project Methodology) in terms of your project's goals, stakeholder involvement, applied research, decision making and reporting.

Goals refer to identified authentic immediate and long term issues that you work towards finding appropriate solutions for. Whilst defining your goals you explore the context and environment of your project and you make the necessary business, sustainable and ethical considerations.

Stakeholder involvement implies that you involve relevant and competent partners in your project from beginning to the end. During your project you communicate constructively with all your stakeholders.

Applied research implies that you effectively use research strategies and methods, like those in the DOT-framework, for your domain understanding and other research activities.

Decision making means that you correctly identify the need for further iterations, using evaluation models like the TIC-tool and your stakeholder feedback.

Reporting refers to well-structured and well-motivated, correct and relevant documents, using APA style referencing for used external sources, as well as using visualizations, concluding your project proportionally covering all four phases of the methodology.

Second Evaluation: week 10

Before I did not put a lot of effort into referencing the sources of my information, now I am trying to implement the APA style referencing, which is an industry standard. While doing so I am also expanding the project proposal and my understanding of the topic that is recommendation systems.

I am also asking of my client to be critical of the things that are happening with the project and their opinion.

Not only that but when it comes to the group project we as a whole, started to separate and use the tasks that we are going to do in Jira, which is also way of organizing the work, that is widely used. When it comes to the tasks and how we are going to be working on them we also started implementing git and have separated the pieces of data that each person is going to work on, as well as the models and have set certain deadlines.

Self assessment: Beginning

As a group we have started using technologies that are widely used in the workplace and also started to properly split the work and assign deadlines. When it comes to the individual improvements, I

started looking into a better way of crediting and more sources of information, for the individual challenge.

First Evaluation: week 5

For the goal of this semester, I decided to create a recommendation system, that is about showing video games with similar features, to a customer, who just chose a product previously. The creation of the project follows the industry standard methodological approach “AI Project Methodology”.

I believe it is the best possible approach, since it gives us a structured plan, while also giving us a lot of needed details and versatility that can be used for the creation of the project.

I invited one of my stakeholders to be a friend of mine, who used to just like me play video games. After a small discussion with him we decided to take the approach of creating a content-based filtering system.

For the project I did decide to get a dataset from “Kaggle” to try and create the system. I filtered through it, looked into it’s columns, etc. At first, I thought that it is a great idea to use it for my project, until I understood that the ratings column that I had planned to use was actually missing a lot of data. For now I believe that I can work around the problem

Self assessment: Orienting

I am following the structure of the standard that we need to follow this semester which is “AI Project Methodology”, while also making contact with the stakeholders to see if they are satisfied the progress of the project. The research that I have made has also been shown in the Project Proposal, where everything is referenced, in a way that I believe to be proper.

5. Learning outcome 5: Personal Leadership

Clarification:

You are aware of your strengths and pitfalls in ICT as well as your personal development. To nurture personal growth, you are able to engage in actions that align with your core values, in a way that suits you.

Being aware of your strengths and pitfalls means that you are able to recognize (among other things through self-reflection and asking for feedback) what you are already good at and where growth is still possible.

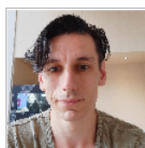
Being able to engage in actions means that you take responsibility, growing towards a professional ICT practitioner, seeing and seizing opportunities in a structured, planned and efficient way.

In a way that suits you means that in the activities you undertake you apply an approach that fits your style of acquiring knowledge and skills.

Second Evaluation: week 10

Our lectures have 'finished', so now we work alone. For that matter I am incorporating routine that will help me stay on track when it comes to coding. If not coding, something else like documentation should be done, but one activity that is connected to the advancement of my individual project should be done a day. This way I make sure to not lose the 'motion' of my project

Taking into consideration the negative thoughts about my work, that are addressed in this feedback:



Hans
Konings

22 March 2024 | [Personal Development Report Evaluation 1](#)

Hi Mihail, as discussed the template has all the crucial ingredients for the PDR and it's good readable, so you should be fine so far. Just write the reflection and growth you got from the feedback and use the other general feedback you've seen in the meeting (page numbering, table of content), though it's just minor things. We see more quality in your work than you do, so you may be more confident in your own work! 😊

Routines are something that I use quite a lot. They have helped me quite a lot and I am also a big fan of 'to-do' and organisation applications. With their implementation, it can help me form habits that will help me stay on track and have a bit of a more positive attitude towards my work, since I start 'accumulating small wins'.

Self assessment: Beginning

I am doing things that I consider one way or another enjoyable that will also help me structure my thoughts and give me a better direction of what to be doing during the day, so that I am on track.

First Evaluation: week 5

As somebody who just came from software, I only have a very rough Idea on how to work with data and how to look into it this semester. It is a crucial point in my learning and I need to get better at it, so what I am doing to work over this problem is be in the lectures and also do some courses that can help me better understand what is done, why it is being done and also, when?

I didn't get a lot of feedbacks last semester, so this one I am trying to have at least one talk with a teacher per week and I also started writing my code in the beginning, not when we could say that it is 'too late'. I started to

I am also trying my best to not go against constructive criticism and just understand more or less their point of view and implement what I have been given as feedback.

Self assessment: Orienting

I am trying to implement different approaches to counteract my previous mistakes and my current lack of knowledge about data.

6. Learning outcome 6: Internship preparation

Clarification:

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 specialization.

First Evaluation: week 10

I went to the career day that was organized to look into the different companies, what they offer and more importantly, what they are looking for. Not only that but I managed to find a site that has some templates for a cv and created one. Also, I added some things to my LinkedIn profile, to the point that I believe it is quite satisfactory.

Self assessment: Beginning

Started looking more into the companies, what they are interested in and what I have to offer. Also, I have updated my LinkedIn profile to a state which I find satisfactory for now and shows what I know.

First Evaluation: week 5

I created my LinkedIn account which is a sure step to help me out. I have started working on how to make my account seem 'reliable', while also thinking about how to make cv. For now I have looked into the templates of 'Microsoft Word' and looked into what a developer who is looking for an internship has in their CV.

Self assessment: Orienting

Making my first steps into the working environment and understanding what I should prepare for and look into