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.

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 | 1 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.

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 epalin very well the domain understanding with diagram but analytic approach need more description and add your research question as well with explaination.

<u>Self assessment: Orienting</u>

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.

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 fro 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.