PERSONAL DEVELOPMENT REPORT

OTHERMEAL

EVALDAS DRASUTIS

Contents

PDR Learning Outcomes

1 - Societal Impact - The student must be able to approach the context and impact of their own AI project(s) from different perspectives in a sustainable way. In addition, the student can reflect on their own choices, considering data legislation and the (possible) impact on society.

2 - Investigative Problem Solving - The student must be able to critically look at their own AI project(s) from different perspectives, recognize problems and come up with appropriate solutions.

3 - Data Preparation - The student must be able to collect data and estimate its quality and usability. The student is also able to adjust the data if necessary for proper usage in their project(s).

4 - Machine Teaching - The student must be able to use data to train models in a way that fits the intended purpose. The student is also able to test whether the models have been adequately trained

5 - Data Visualization - The student must be able to use data to create an interesting, informative, and compelling story in an (interactive) data visualization product, tailored to the right target group.

6 – Reporting - The student must be able to report in a methodologically sound manner on (the outcome of) their own AI projects (project proposal, process documentation, reporting of final results, etc.).

7 - Personal Leadership - The student shows an entrepreneurial mindset regarding their own AI project(s) and personal development, while being aware of their own learning capacity and keeping in mind professional ambitions in their future work field.

8 - Personal Goal - With this learning outcome, the student can set its own goal in relation to their future field of work. (Describe this Learning Outcome in this PDR).

		Advanced	Proficient	Beginning	Orienting	Undefined
1	Societal Impact					
2	Investigative Problem Solving					
3	Data Preparation					
4	Machine Teaching					
5	Data Visualization					
6	Reporting					
7	Personal Leadership					
8	Personal Goal					

1. Learning outcome – Societal Impact

In this chapter, I will summarize the research process based upon projects and challenges I participated in. Defining the societal impact of AI, its ethics and its different biases which are by-products of the current and future technologies.

1.1. First evaluation

Achieved level – week 3: Undefined

During the introduction to societal impact from the workshops I got introduced to a new point of view on possible unintended consequences of current and future technologies. Considering the given new perspective, I tried to apply the same base questions to my project idea to foresee possible societal issues. So far, I need to research more upon the topic to gain more insight.

For this evaluation phase, I believe I have reached the level of 'Undefined.

1.2. Second evaluation

Achieved level – week 7: Orienting

During this iteration in my study, I had an opportunity to participate in several workshops regarding positive and negative AI effects on society. I got informed on the AI capabilities that could improve daily human tasks and improve human productivity, although AI could have coded biases that could be missed very easily or entirely ignored due to a presumption that the build product is doing what it was designed for purely because it is using AI.

Some research I have done based on this subject can be retrieved in the Bibliography (Page…)
- (Shane, 2019)
- (Marr, 2021)

Because I have researched on the positive and negative impact of technology in a society and its vulnerabilities as well as the effects it could cause for people, I believe that I have reached the level of 'Orienting'.

1.3. Third evaluation

Achieved level – week 10: Beginning

During this iteration I had partaken in several workshops regarding AI unintended behavior affecting people's life immensely through coded bias in software. One of the most thought-provoking parts to my research was watching a recommended documentary by the teachers called 'Coded biases' as well watching ted talks from the creator of the movie "Joy Buolamwini". It made me aware of what problems the current tech is facing and what future problems could arise due to the lack of universal user testing.

In addition, another important question that was raised is that the developed AI that is widely used can be harmful to some groups of people primarily people of color. As some people might not get any health benefits, housing, or loans because the application prioritizes white clients over other ones presuming from the existent data. (Buolamwini, 2017)

For the personal project, I tried considering any negative factors that could apply to my project taking in consideration the given examples from the workshops. The only aspect I could see that being applied in my software is that my application could be ambiguous in how it accurately displays nutritional data where it could negatively affect someone who might need an exact amount of nutrients in a certain recipe.

For participating in workshops and challenging myself to assess my own project from a critical perspective I believe I have achieved the level of "Beginning"

1.4.   Fourth evaluation

Achieved level – week 17: Proficient

During the final iteration I have written several documents regarding my progress with machine learning algorithms and models used in the application as well wrote a document addressing societal impacts that could be applicable to the developed software. Through the progress of writing the documentation on societal impacts I was able to critically assess my software depicting features that would need improvement in future versions. The improvements could benefit different social groups by including more inclusive determination of recipes origin from the recipes contained ingredients.

By documenting and researching possible shortcomings of the software I believe I have achieved the level of "Proficient"

2. Learning outcome – Investigative Problem Solving

In this chapter, I will approach different problems and try to define the research questions applicable to the appropriate problems. I will also support my own choices with critical analyses while searching for any alternative perspectives.

2.1.   First evaluation

Achieved level – week 3: Undefined

During this iteration I have defined what my personal project idea could be and what societal factors could be considered useful for society. In addition, I must consider what data I must use for delivering a complete description of my project's functionality. Furthermore, I was working on the group project solution defining the idea for the air quality company.

For the current reporting and documentation process I went through, I believe I have reached the level of 'Undefined'.

2.2.   Second evaluation

Achieved level – week 7: Orienting

For the individual project I have created an initial project proposal where I analyzed my idea to describe the requirement, deliverables, scope, benefits, and its goal which can be found in the git repository. In addition, I have made a wireframe to make a graphical template visualizing the possible interfaces for the application.

Through workshops I got some information upon identifying 'Key point index' values on different datasets which allowed me to inspect my datasets critically. Allowing me to separate required values to get critical data defining my datasets status. This distinction allowed me to define average amount of reviews that exist on individual recipes. Providing us an average estimate of user interactions on recipes that could be used later, on applying a machine learning model.

As for the group project, I have participated in the group sessions where we analyzed the client's proposal to formulate the initial projects idea. By Defining the projects requirement, deliverables, scope, benefits, and its goal we expanded on the idea taking in consideration the

client's requirements. By researching we formulated the main research questions regarding the building requirements criteria, for determining the most suitable building site. From these research questions I have tackled selection of data sets, the creation of functional requirements, and the wireframe development which was a collaborative activity.

For the current reporting and documentation process I went through, I believe I have reached the level of 'Orienting'.

2.3. Third evaluation

Achieved level – week 10: Beginning

For personal project I developed a front end for my application that I could display the datasets contents and improve upon my web development skills. For this I used react to allow faster interface development. In addition, I have defined my datasets contents merging them and adding positive and negative ingredient preference on the tables. This will allow me to train the data sets' ingredients based on a person's positive and negative inputs on individual ingredients.

For the group project I participated in the meetings discussing further development of the features and additional research that is required to be done. Additionally, we worked on delivering application components for creating the prototype of the application. I had a chance to create the metrics page using "ChartJs[i]" to create various graphs that could be used for showcasing the estimate chemical emission in certain areas and having an overview of different types of chemicals that are detrimental to the environment, which people should be cautious of.

For the current progression on my personal project and work progression on the group project's application development I believe I have reached the level of "Beginning".

2.4. Fourth evaluation

Achieved level – week 17: Beginning

For the personal project I have trained a KNN and Decision tree and random forest models to recommend the user a recipe that is high in interactions ratings and recommends recipes based on ingredient preference. By evaluating unlabeled recipes which had none of the liked or disliked ingredients I was able to determine an affinity index by checking reoccurrence of ingredients in already liked ones. This was achieved by exploring spatial KDtree to determine the similarity in the encoded recipe ingredients. This allowed me to get accurate representations of recipes that the user might like or dislike.

To be able to recommendation based on multiple inputs I looked up in multi-dimensional arrays. As the data that I was using for the recipes consisted of dense amount of categorical data that was not applicable to encoding as a separate column.

For the project I have developed a recommendation system based on cosines similarity giving recommendations on searched recipe names.

For the current progress with recommendation

3. Learning outcome – Data Preparation

In this chapter, I will describe my work cleaning, filtering, and preparing the initial data using varied methods to highlight significant data in my models to be later used for further implementations for my projects.

3.1. First evaluation

Achieved level – week 3: Undefined

From the workshops I got introduced to a data preparation process. From this I was able to search for a recipe data set that would be extensive enough to be able to be used as usable data for processing. But regarding the data cleaning and quality estimation, I must do some more research.

For this reason, I believe that I have reached the level of 'Undefined.

3.2. Second evaluation

Achieved level – week 7: Beginning

During the group work process, we worked together on finding datasets that would add value to the product's functionality. This allowed us to debate on its usefulness and choose the most significant data that would give value to a product. This documentation could be found on teams file folders.

To better my understanding on Data preparation I have followed several tutorials that teach how to use the Pandas and NumPy libraries which allow me to manipulate the datasets.

I have looked through the data set and defined what values I identify as key values. To get a numerical indication for later machine learning. I have tried getting all the unique ingredients to represent possible selections in the recipes.

I started manipulating the data set first extracting the unique ingredients that exist in the data set and getting a count of them to get an overall overview of the existent data.

Because of the following research, I believe I have reached the level of 'Beginning'.

3.3. Third evaluation

Achieved level – week 10: Beginning

During this iteration I have worked on merging the recipe table with the user interaction table giving the recipes both the average rating and users preference indicator. This allowed me to have more versed data on the item to determine users' affinity to it. In addition, with received values I tried to plot it out to see the usability status of the data.

For the group project I have been working on displaying statistics on air quality around Eindhoven. Using the TNO air data set we searched for the most crucial chemical types for people which is based upon the research questions from the previous iteration. Which helped us determine what data, we would like to use for displaying statistical overviews of the chemical pollution in the area.

Because of the following research I believe I have reached the level of "Orienting".

3.4. Fourth evaluation

Achieved level – week 17: Beginning

During this final iteration I have worked on separating my notebooks to extract the modified data that could be used for the recommendation system. This separation allowed me to notice that some of the data that I was using from the start was missing over 53,000 rows of data when merging different data-frames. To which I had to clean up the data-frame again.

In addition, I tried adding nutritional data to the recipe data set to comply more with societal impact effect to the software.

Because of the following data preparation, I believe I have reached the level of "Beginning".

4. Learning outcome – Machine Teaching

In this chapter, I will report the use of the acquired data to modify and prepare the information to be used for AI training models providing an effective solution to my problems.

4.1.   First evaluation

Achieved level – week 3: Undefined

I have gotten an introduction to the machine learning in the workshops. So far, I still need to figure out a way on how I would apply AI to recognize the user's preferences both in saved ingredients and recipe selections.

For researching requirements for my project, I believe that I have reached the level of 'Undefined'.

4.2.   Second evaluation

Achieved level – week 7: Orienting

From the given workshops on machine learning I used the provided notebook tutorial to better my understanding of the KNN algorithm and decision tree models. From the given notebooks I was able to further determine what machine learning models would suit my project more.

As I will need to do recommendations, I have considered 2 types of machine learning models. First being the Linear regression model to make recommendations based on preferred ingredients. Second being the K-nearest neighbors algorithm model to get recipe recommendations based on similar user patterns. But to apply it I need to prepare my data and research more on the linear regression model.

Because I have trained an AI model from the notebook and started researching on appropriate models for the personal project, I believe I am at the level of "Orienting".

4.3.   Third evaluation

Achieved level – week 10: Beginning

I have been working on my personal project trying to apply the linear regression model to be able to predict users' preferences based on ingredients that the user likes and dislikes. This allows the recipes in the data set be marked true with a value of '1' if the recipe ingredients exist with these values or if it contains ingredients disliked marking it false with value '-1', and if neither apply the value will maintain being 0 as undefined. In addition, I have tried applying the KNN algorithm to the user rating average to group up possible user interactions.

Because I have been working on my personal project machine learning training and data modification, I believe I am at the level of "Beginning"

4.4.   Fourth evaluation
Achieved level – week 17: Beginning

I have worked on improving my KNN and decision tree models to provide a better result score. As well removing unnecessary models that I used within the workbook. I mostly focused on recommending recipes with several ingredients but to no success. I mostly tried using several models to predict different features.

Because of the following progress with developing a hybrid recommender system, I believe I have reached the level of "Beginning".

5. Learning outcome – Data Visualization
In this section, I will describe the process of preparing and processing the data to be visualized in insightful graphs and charts giving indications of data quality.

5.1.   First evaluation
Achieved level – week 3: Undefined

I still have not prepared any data sets to visualize the obtained data.

Because of this I believe my learning outcome is "Undefined".

5.2.   Second evaluation
Achieved level – week 7: Orienting

In the personal project I have researched the math plot library which allows use of data visualization with graphs. I have practiced recreating the plots using the university notebooks as a guideline, this allowed me to visualize the overview of the current data status. I have made a bar plot to define the unique ingredients and a scatter plot to visualize the recipe ratings. Adversely the returned graphs showed a lot of empty values resulting in not properly scaled data. Furthermore, I will have to access the data to provide a clear definition for the plots.

Because I have reviewed the notes from the university and tried to plot the data using various graphs from the provided info as well as tried to apply the graphs to my personal project by visualizing the existent data, I believe I have reached the level of "Orienting".

5.3.   Third evaluation
Achieved level – week 10: Orienting

During this iteration I have been working on the group project to deliver a statistics section of the dashboard. I have created a few template charts such as scatter, bar, line, and doughnut charts to be able to display unique aspects of the data.

For the individual project I have been continuously creating different plots to display my data to get more insight on its properties. But after the feedback session I was informed that I need to rework my graphs because they need rescaling to normalize the data.

For this I believe I have achieved the level of "Orienting"

5.4.   Fourth evaluation

For the individual project I have rescaled my data to display effective plots describing the data status. By removing several outliers, I was able to scale the plots to have a more accurate representation. By plotting several heatmaps to overview my similarities between my data-frames I noticed that the accuracy was low in-between the components. Making me believe that I still had too little features in my datasets.

As well plotted different machine learning models scoring to determine which one provides me with most accurate predictions.

Plotted important features from the data-frame to show overviewed data status and amounts of existing outliers.

6. Learning outcome – Reporting

In this section I will report my professional attitude on how I interact with the consultants on my personal development progress. Providing my actions both in group and personal project and the application of received feedback.

6.1.   First evaluation

Achieved level – week 3: Undefined

For the first iteration I have yet to set my first meeting with a consultant because of little understanding of the topic and need of further research in both personal and group project.

Because of this I believe I am at the level of "Undefined".

6.2.   Second evaluation

Achieved level – week 7: Beginning

I have worked on documenting my personal project proposal, PDR, and group projects proposal where I got to document the required features. This allowed me to create an estimation on what I will have to deliver in the following iterations. With having an idea on how I will proceed on to the next steps I made my first meeting with a consultant giving an overview of my progress, showcasing a wireframe of the application for reference and the following steps I'm considering of taking. The given feedback provided me with the steps that the consultant expects me to follow for the next iteration.

Because I have worked on documenting the projects and researching the topics in-depth and asking for feedback from consultants, I believe I at the level of "Beginning".

6.3.   Third evaluation

Achieved level – week 10: Orienting

For this iteration I have been working on developing my machine learning algorithms that would help me filter the recipes based on liked ingredients as the teacher recommended. In addition, I have created the front end for my personal project that I could showcase the applications interface. By giving an overview of my current progress the consultant recommended me to search for some Kaggle examples online on notebooks to better understand how I can apply machine learning to the current data that I have as well as understand how to scale and normalize the existent graphs. For further work I'm expected to train my data and give a true or false estimate for the recipes that the user would like.

For the interactive attitude with the consultants, I believe I reached the level of "Orienting".

6.4.   Fourth evaluation
Achieved level – week 17: Proficient

For the final iteration I have shown different consultants my personal project progress explaining my problems that I was facing that are hindering me from generating wanted recipe recommendations and the current work that I have produced so far. As well detailing the different choices I took in applying machine learning to my project. I got recommended to investigate multi-categorical vectorization which would help with categorizing the models features.

For the interactive attitude and applied feedback, I believe I reached the level of "Orienting"

7. Learning outcome – Personal Leadership
In this section I show my entrepreneurial mindset approach providing a description on personal development both in communicative approach and independent work done in my personal and group projects.

7.1.   First evaluation
Achieved level – week 3: Undefined

For this iteration I must still set up my project plan that I could follow along, this will require more research of the project's idea. This documentation is required for smoother development in later stages of project implementation.

For this I believe I'm still at the level of "Undefined".

7.2.   Second evaluation
Achieved level – week 7: Orienting

During the group project's meetings, I had discussed my ideas on the project's initial idea and tried to understand what the essential functionalities of the application are. In addition, I had an opportunity to question the client on dataset quality and restrictions of building permissions. This allowed us to expand the research questions that we would need to adhere to for later when we are developing a prototype.

For personal project I have developed a project proposal where I described the required features that I foresee implementing. Defining the required features in a MOSCOW prioritization table to arrange the important features I want to deliver in the future iterations. To help me visualize what I believe my website should have I made a graphical wireframe displaying the website pages.

Feedback is detailed in **Figure 2-3** in section feedback references

Because of active behavior in group meetings, personal project documentation, and feature planning organization I believe I am at the level of "Orienting".

7.3.   Third evaluation
Achieved level – week 10: Beginning

For the group project I have been discussing the further implementation of the application and additional requirements needed from the given datasets from the client. In addition, I worked with the group creating the posters for the midterm delivery. Summarizing the main functionalities and raising the research questions to define the outline of the projects purpose.

For the personal project I have been mainly focusing on researching and applying different solutions to better my machine learning and data preparation. I have been following the consultants' advice on what type of recommendation I should prioritize and what later steps I could take using the KNN algorithm. Furthermore, I'm going to work on training my recipe data set using liner regression to give estimations of still neutral data existent in the table.

Feedback is detailed in **Figure 4-5-6** in section feedback references

For working on group project's poster and participating in the poster fair as well working on personal deadlines I believe I reached the level of "Beginning".

7.4.   Fourth evaluation

For the personal project I have been mainly working with machine learning which took too much time. I was able to get results on possible predictions on recipes based on ingredient name. As well recommending ingredients based on other user interactions. This hindered my expected goals for my project progress because I was not able to predict recipes based on multiple ingredients. Because the dataset contained too many categorical items to be used with hot encoding I had to resort to other solutions.

Feedback is detailed in **Figure 7** in section feedback references

For working on machine learning as well vectorizing data, I believe I reached the level of "Orienting"

8. Learning outcome – Personal Goal

In this section, I will highlight my personal goals for the iterations which is based on my perspective of what is important to grow professionally.

8.1.   First evaluation

Achieved level – week 3: Undefined

For the personal goal I must set a clear goal on what I want to improve upon. Although as an initial idea I want to learn how AI works, what machine learning model types are, and how can I apply it in practice for my personal project.

For the moment I still need to research more on machine learning to define my goal for this. I believe I am still at the level of "Undefined".

8.2.   Second evaluation

Achieved level – week 7: Orienting

For this iteration I participated in the lecture that give different aspects of AI appliance. I got to learn societal impacts of AI and multiple machine learning models as KNN and decision trees which allow the data to be trained to make predictions based on scaled inputs.

Personally, I feel I need to research more about AI models and the different uses of them for my personal project, so that I can be able to apply the best machine learning model types which in turn would make the best predictions.

Because I have participated in workshops regarding the use of AI and researched online via video tutorials showcasing AI application, I believe I have reached the level of "Orienting".

8.3.   Third evaluation

Achieved level – week 10: Beginning
For this iteration I have been researching more in-depth about machine learning algorithms to increase my understanding of it and its application in my project.
In addition, I have worked on creating the front end to the application allowing me to display the filtered content but also improve my web design skills.
For the group project I continued working on the metrics page where I still need to scale the chemical data to have an overview estimate on existent data

Feedback is detailed in **Figure 5-6** in section feedback references

For further development in the personal and group projects I believe I have reached the level of "Beginning"

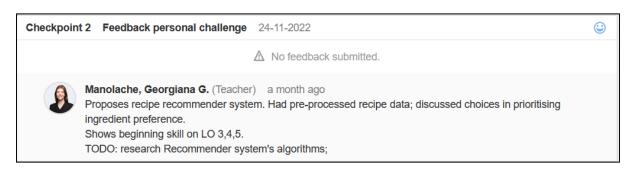## 8.4.   Fourth evaluation

Achieved level – week 17: Beginning
For this iteration I have achieved in applying several machine learning models to my software which allowed me to recommend a recipe based on recipes ratings as well on recipes name. Due to working on the machine learning I have spent less time improving my web design to display the items through an API.

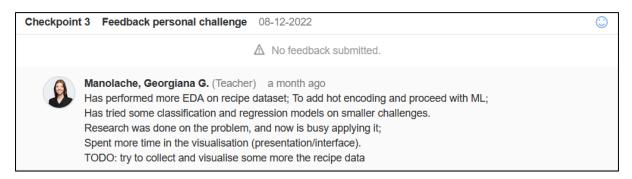For further development in the personal and group projects I believe I have reached the level of "Beginning"

Feedback references

FIGURE 1

FIGURE 2

FIGURE 3

FIGURE 4

Checkpoint 5   Feedback personal challenge ML & EDA   15-12-2022

⚠ No feedback submitted.

FIGURE 5

Checkpoint 6   **Feedback personal project**   22-12-2022

Drąsutis, Evaldas E.   14 days ago
Discussed on required expansions to the inputs features. Suggested additional research on recommendation systems. In addition, remove unneeded classification that does not benefit the data prediction.Lastly, write a report regarding used machine learning algorithms in my project.

FIGURE 6

Checkpoint 7   **Friday 23 Dec - Personal Challenge (Martijn)**   23-12-2022   ☺

⚠ No feedback submitted.

Lamers, Martijn M.J.M. (Teacher)   14 days ago
Hi Evaldas,

PS. You were absent today in building TQ5 for our appointment, unfortunately, but I agreed in talking to each other ONLINE then ...

You discussed and demonstrated your individual challenge on a recipe recommendation device to me.

At first, it seems the algorithm should have to predict 1 particular recipe, based on the ingredients that one particular person likes ... Right? So that's like a categorization problem for a multi-category situation.

You used indeed a random forest / decision trees algorithm for that. And looked at KNN.

But you also showed me a regression approach ... What does a certain number mean then in your context?

You showed various graphs, to investigate the (effect of) possible outliers. For example, if the recipe preparation time is too long, it shouldn't be considered for recommendation.

I think you choose a quite complex project with many ingredients as input and many possible recipes as output. Because of that it's a bit hard to see how well you understand the "basic ML concepts" ... But you did a lot, to make sure the recipe recommendation would work. Chapeau!

You are currently trying to combine your 2 different Jupyter notebooks. And also to improve the accuracy of your machine learning model(s) for the recipe recommendation.

I would to see your final version in the first week after the holiday in January 2023.

Have a nice Christmas holiday and end of the year!

Regards,
Martijn

FIGURE 7

Bibliography

Buolamwini, J. (2017, March 29). *How I'm fighting bias in algorithms*. Retrieved November 27, 2022, from Youtubr: https://www.youtube.com/watch?app=desktop&v=UG_X_7g63rY

Marr, B. (2021, January 15). *What Is The Impact Of Artificial Intelligence (AI) On Society?* Retrieved
    November 27, 2022, from Youtube:
    https://www.youtube.com/watch?app=desktop&v=_eaJd3muf2E

Shane, J. (2019, November 14). *The danger of AI is weirder than you think*. Retrieved November 27,
    2022, from Youtube: https://www.youtube.com/watch?app=desktop&v=OhCzX0iLnOc

---

[i] Graph library