**PDR Document**

**Landscape classification**



**Name:** Hristo Hristov

**Email:** [h.hristov-ab@student.fontys.nl](mailto:h.hristov-ab@student.fontys.nl)

**Student Number:** 4170431

**Teachers:**

Zhao Qin – [q.zhao@fontys.nl](mailto:q.zhao@fontys.nl)

Hans Konings – [h.konings@fontys.nl](mailto:h.konings@fontys.nl)

John Litsenburg – [j.litsenburg@fontys.nl](mailto:j.litsenburg@fontys.nl)

**Fontys University of Applied Sciences**

**Eindhoven – Netherlands**

**2022**

Contents

[Introduction 3](#_Toc101400062)

[Revision table 3](#_Toc101400063)

[Learning Outcomes 4](#_Toc101400064)

[1. Data preparation 4](#_Toc101400065)

[2. Data analysis & model engineering 7](#_Toc101400066)

[3. Reliability and transparency 12](#_Toc101400067)

[4. Targeted interaction 14](#_Toc101400068)

[5. Future orientation 19](#_Toc101400069)

[6. Investigative problem solving 21](#_Toc101400070)

[7. Personal leadership 24](#_Toc101400071)

[8. OE4 Internship Preparation 28](#_Toc101400072)

# Introduction

This document is written with the purpose of documenting my personal development through the semester 4 of AI42. A little background on me: I come from Bulgaria, I started my study at Fontys in 2019, I chose to pursue Software as my specialization because while doing the orienting phase that subject struck me as the most interesting. I come to this semester hoping that I will learn new technologies that will not only be helpful to me in my future career but will also let me create something fun that has the ability to be useful.

|  |  |  |
| --- | --- | --- |
| Revision table | | |
| Doc.No | Date | Change |
| 0 | 3/10/2022 | Document created |
| 1 | 4/20/2022 | Added additional proof, updated self-evaluation, added internship preparation. |

# Learning Outcomes

## Data preparation

* 1. Definition

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:

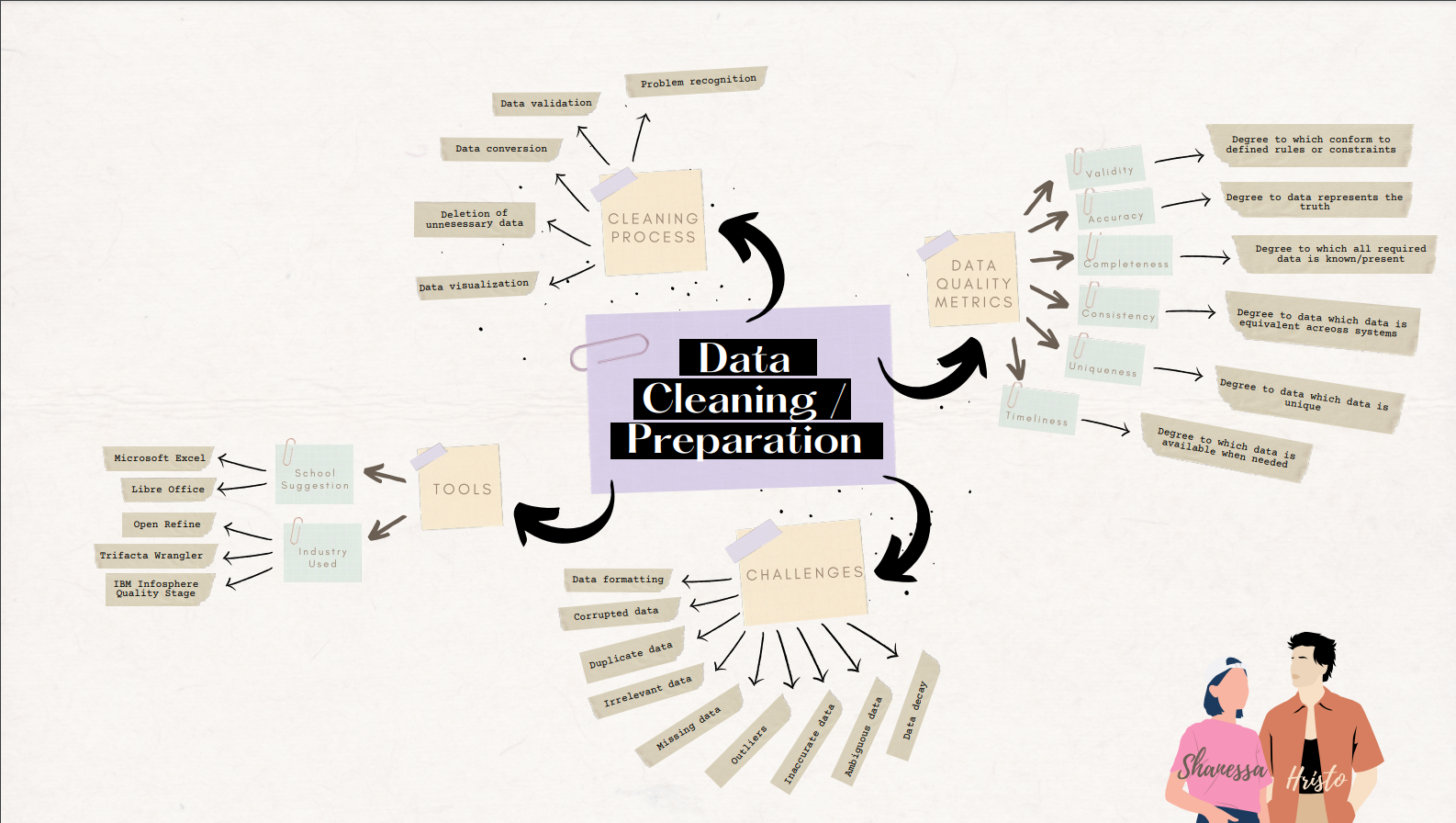
Preparing a dataset consists of extending it with additional data and cleaning the data according to theories of data quality, in such a way that the process of cleaning and preparing those data is repeatable, transparent to others, and the results are suitable for data analysis and/or modelling. Storing a dataset includes investigating cloud solutions and arguing whether they should be used.

* 1. Self-evaluation

| Advanced | Proficient | Beginning | Orienting | Undefined | **Total points** |
| --- | --- | --- | --- | --- | --- |
| 4 Points | 3 Points | 2 Points | 1 Points | 0 Points | 4 Points |

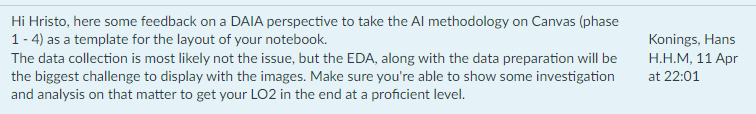
|  |
| --- |
| Orienting |
| I believe I am at the Orienting level because I have gotten familiar with the 5 different data quality measures: Validity, Accuracy, Completeness, Consistency and Uniformity. I have also started diving into the different techniques used for data cleaning (such as dropping data, filling in missing data, hot-shot encoding). I also have done a research exercise regarding data cleaning. |
| Beginning |
| I believe that I am in beginning because in regards to my individual project I have begun my data cleaning process to make sure that all of the data that I will be using is not corrupted. While loading each image into my notebook I also make sure to resize it so that our model can more quickly process each of my 4000 images. I also change the color of the picture into RGB, just in case that it already is not in that color scheme. |

* 1. Proof

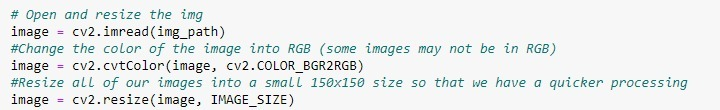


*Explanation: A Mind map I made in regards to the different data cleaning processes/preparation techniques*

**

**

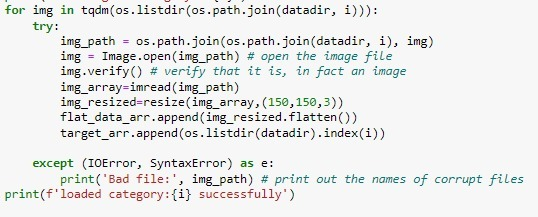
*Explanation: I think that in regards to data analysis there is quite a lot of options for exploring my image data, not only using labels but also the images themselves.*



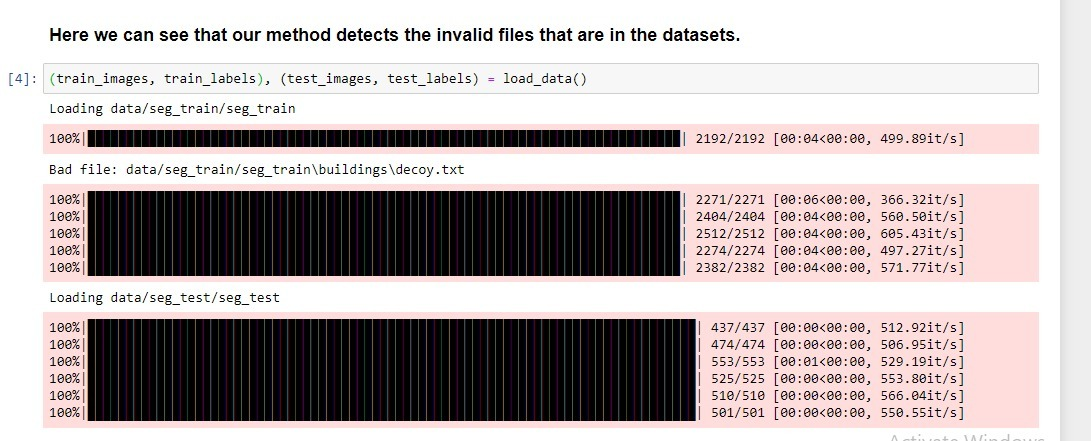
*Explanation: This is a code snipper where I take the images that I have a I resize them to make sure that my model can more quickly process them (as I have over 4000 landscape images) and I change the color of the image to RGM to make sure that all of my images are the same color scheme (in case some images are not).*

|  |  |
| --- | --- |
| Date | Meeting with Hans online through teams |
| 11.04.2022 | Topics discussed |
| * Reasons for my inability to come in person to school * One way to explore my data is by showing the images and their corresponding labels * Since my data is already clean, implementing some sort of check to see if any future data might not be clean or corrupted is a good idea | |

*Explanation: My request for a meeting with Qin and the topics discussed in our meeting*



*Explanation: This line of code checks if each image in our dataset is actually an image and whether or not it is corrupted*



*Explanation: Here we can actually see that our code detected a txt file and did not load it.*

## Data analysis & model engineering

* 1. Definition

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.

* 1. Self-evaluation

| Advanced | Proficient | Beginning | Orienting | Undefined | **Total points** |
| --- | --- | --- | --- | --- | --- |
| 4 Points | 3 Points | 2 Points | 1 Points | 0 Points | 4 Points |

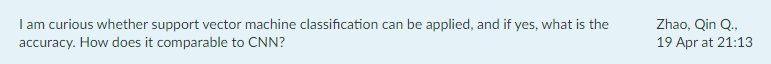
|  |
| --- |
| Orienting |
| I believe I am at orienting level because I have started applying different algorithms (2 classification and 1 regression algorithm) to my data. I have also tuned different hyper-parameters such as KNeighborsClassifier for nearest neighbors and n-estimators for decision trees |
| Beginning |
| I believe I am at beginning because I have begun analyzing my data in regards to my individual project, I have also began experimenting with different models (Keras, SVM) to determine which model best fits my data and gives me the best results. |

* 1. Proof

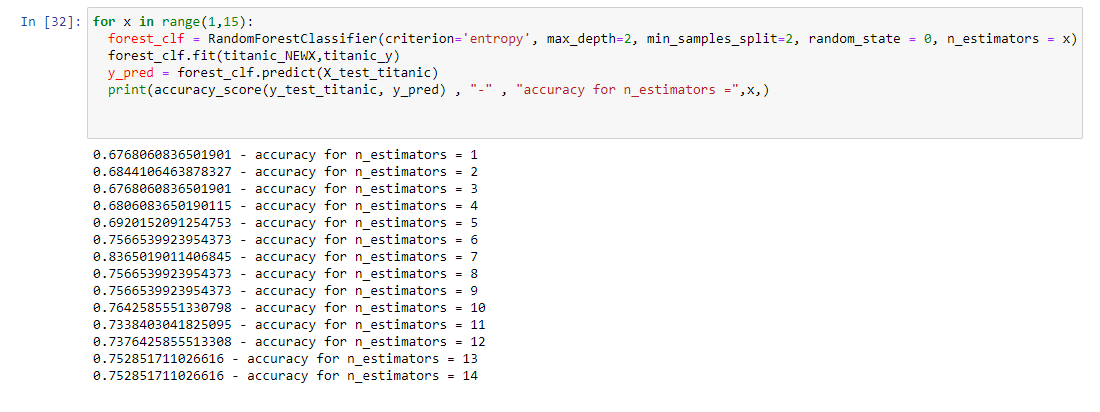


*Reflection: I think I did a good job with the k-nearest neighbor algorithm, however I over fitted the data in some instances of my usage of the k hyperparameter.*

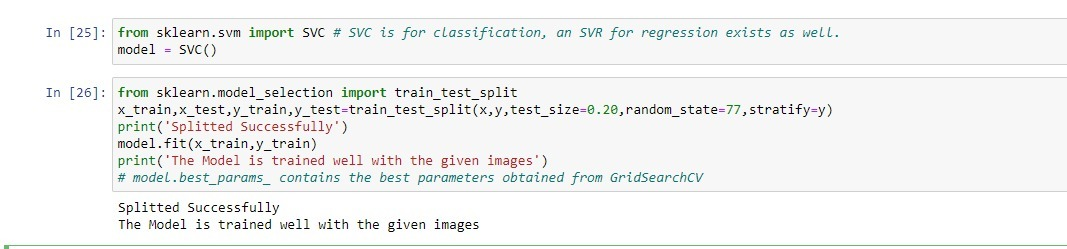
**

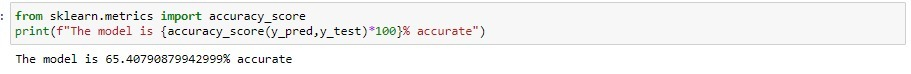
**

*Reflection: I think comparing the 2 different models can be quite interesting but fine tuning the for hyperparameter both models will take quite some time*

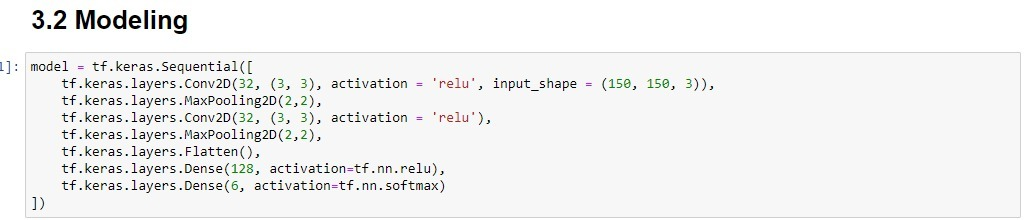


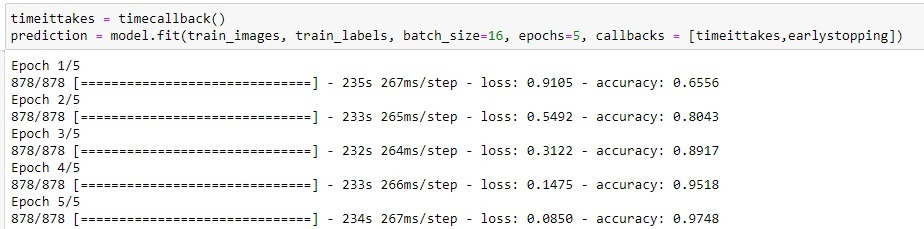
*Explanation: Here I try to use a different k values for my model to see which values gives me the most accuracy.*

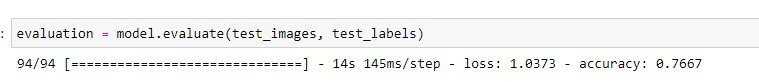




*Explanation: My SVM model with 65% accuracy.*

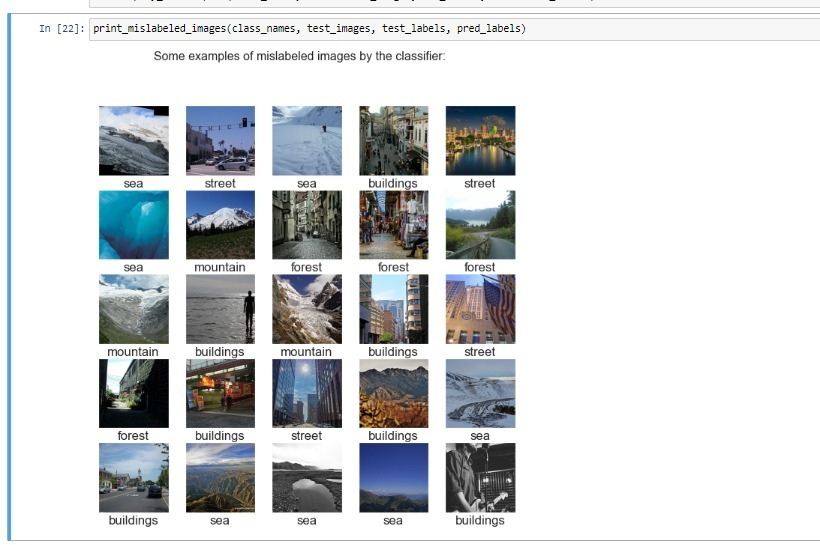




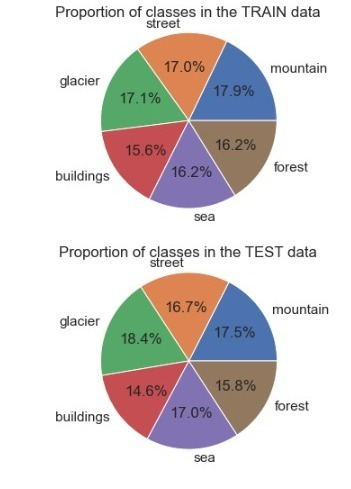


*Explanation: My Keras model with 98% accuracy on my train data and 76% accuracy on my test data.*

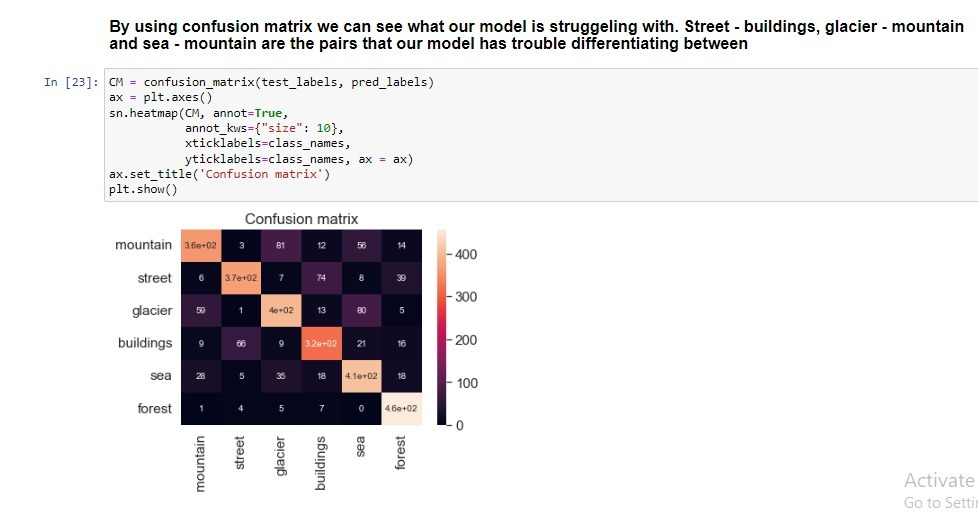
*Overall Explanation: Here I make 2 models and I fit my data into them. When comparing we can see that our Keras model has an advantage over the SVM model.*



*Explanation: Exploration of my dataset*



*Explanation: Analysis on the class distribution in my dataset*



*Explanation: Analysis on what pairs of classes my Keras model is having trouble differentiating between.*

## Reliability and transparency

* 1. Definition

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

Clarification

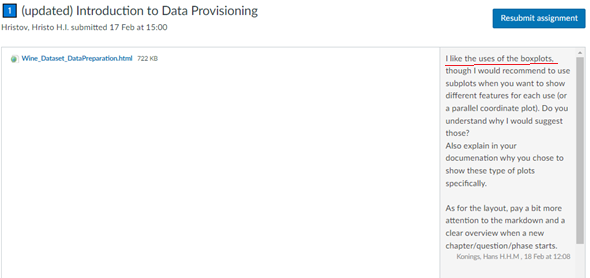
Reliable means that conclusions are supported by methodologically acquired and evaluated results, recommendations make sense in regards to the domain knowledge, and used sources are referenced appropriately. Transparent refers to the process being clear to such a degree that it is reproducible, results being explainable to humans and based on decision making that is considered fair, whilst eliminating bias.

* 1. Self-evaluation

| Advanced | Proficient | Beginning | Orienting | Undefined | **Total points** |
| --- | --- | --- | --- | --- | --- |
| 4 Points | 3 Points | 2 Points | 1 Points | 0 Points | 4 Points |

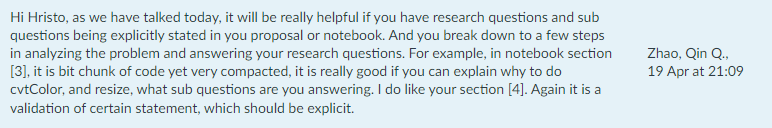
|  |
| --- |
| Orienting |
| In regard to this learning outcome, I believe I am at orienting because I am still learning how to make my notebooks as readable as possible while also providing good explanation in the comments as to why I have done the things I have done. |
| Beginning |
| I believe I am at beginning because in my notebook I make sure that I explain my though process and also my code so that the person reading it can understand the reason I have chosen to solve the problem that way that I have done. |

* 1. Proof

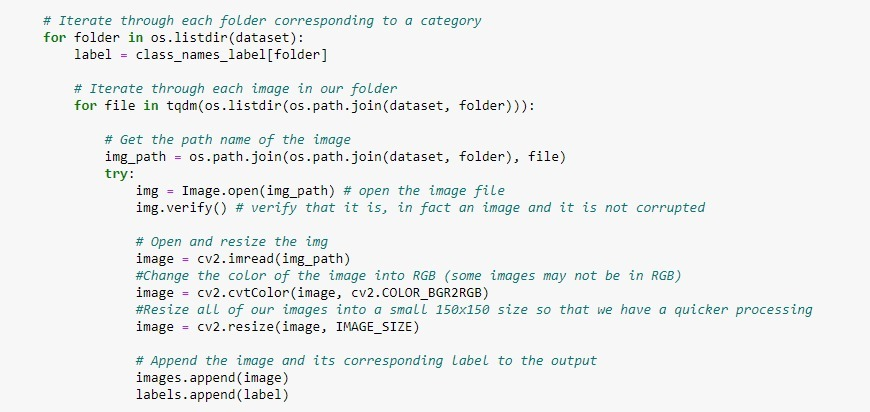


*Reflection: I think here I could have used different plots for visualization the data instead of using mainly boxplots, however even using only boxplots does reveal a lot of information.*

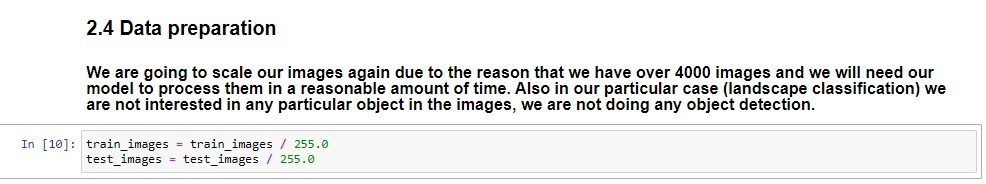
**

**

*Reflection: Here I receive feedback on parts of my notebook where a clearer explanation on why I do certain processing of my images can be of great help. I have definitely taken note and will make sure that every line of code has an explanation behind it and a reason for it.*



*Explanation: In this compact piece of code I make sure that every line of code has an explanation and it can be understood what the code does.*



*Explanation: Here I explain the reason why we need to scale our images.*

## Targeted interaction

* 1. Definition

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.

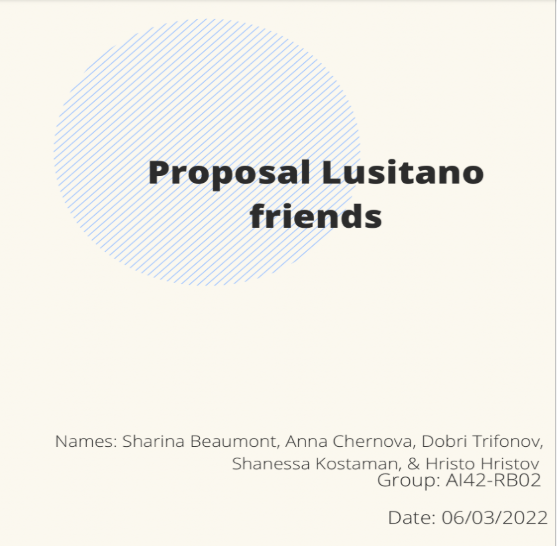
Based on your message, your position and the person(s) you address, you choose the right channel and format to communicate results, including appropriate data visualizations (such as data stories, infographics, or (a set of) static or interactive plot(s)).

* 1. Self-evaluation

| Advanced | Proficient | Beginning | Orienting | Undefined | **Total points** |
| --- | --- | --- | --- | --- | --- |
| 4 Points | 3 Points | 2 Points | 1 Points | 0 Points | 4 Points |

|  |
| --- |
| Orienting |
| I believe I am still at orienting because while I am learning a lot about how to do things such as my project proposal for my individual challenge and data preparation/analysis/modeling for my challenges. I also learning how to properly present my finding from the research I have done. |
| Beginning |
| I believe I am at beginning because as I have started working on my notebook I have also started to communicate more and more with the teachers to make sure that I am on track with my project and to keep them informed of my progress. In regards to the group project I make sure that we as a group keep our stakeholders informed of our progress and we ask appropriate questions to make sure we are fulfilling our client’s requirements and needs. |

* 1. Proof



*Explanation: A mock project proposal that we did for the group project*

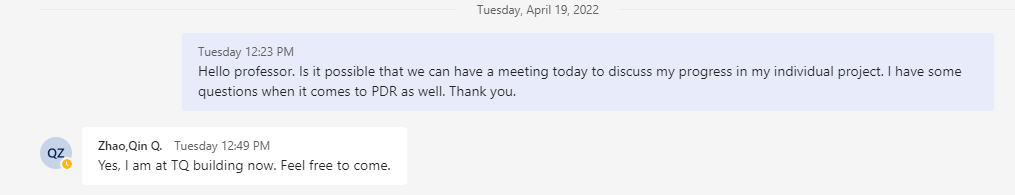


*Explanation: A presentation regarding our proposal for the group project*

Graphical user interface, text, application, email

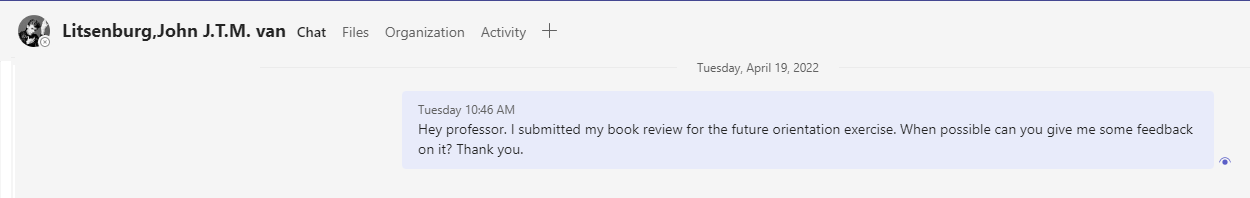
Description automatically generated

*Explanation: An email we sent to our company contact person so that we can clarify some questions that we had.*

**

|  |  |
| --- | --- |
| Date | Meeting with Qin in person in TQ5 building |
| 19.04.2022 | Topics discussed |
| * Improvements to my existing Keras model can be made * More explanations in code snippets are needed * Add some main research questions and answer them so that it is easier to follow my thought process | |

*Explanation: My request for a meeting with Qin and the topics discussed in our meeting*

**

|  |  |
| --- | --- |
| Date | Meeting with John in person in TQ5 building |
| 19.04.2022 | Topics discussed |
| * Different views of people in the A.I. industry * Doomsday preachers / people being let down by the stupid of current A.I., different sides of the views held by people * Difficulty of teaching simple objects/creatures recognition to a machine. * Big figures like Elon Musk joining forces with other philanthropists to answer the question: What is intelligence? | |

*Explanation: My request for a meeting with John and the topics discussed in our meeting*

## Future orientation

* 1. Definition

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

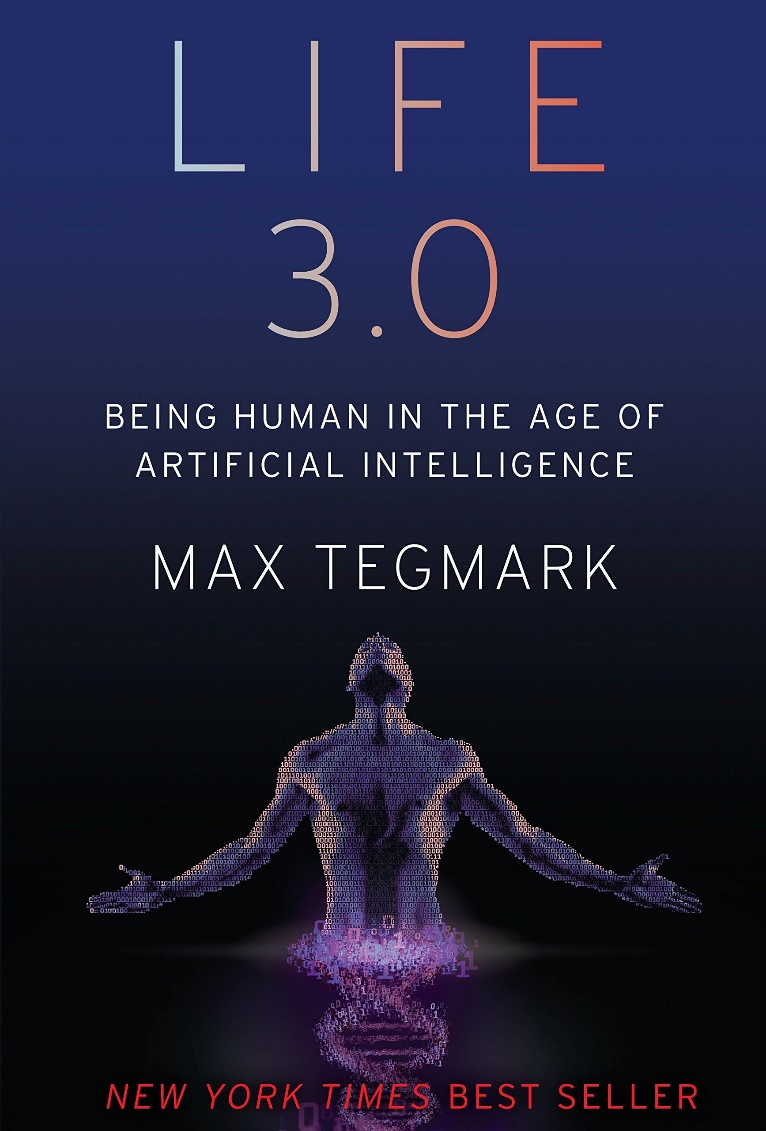
Assess the context means the organizational and societal environment in which the data analysis project takes place. You show that you can identify the hallmarks and roles of the environment of the assignment and have a keen eye for a future-oriented, sustainable embedding of your work in an organization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

* 1. Self-evaluation

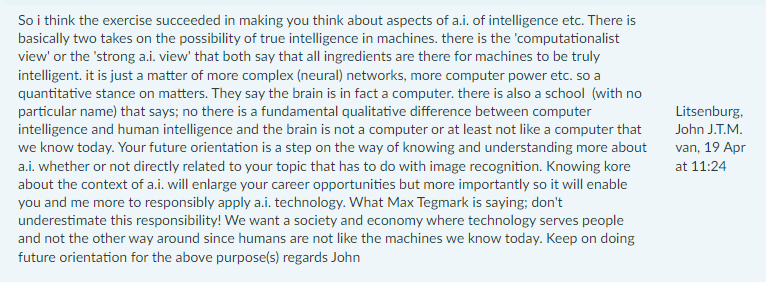
| Advanced | Proficient | Beginning | Orienting | Undefined | **Total points** |
| --- | --- | --- | --- | --- | --- |
| 4 Points | 3 Points | 2 Points | 1 Points | 0 Points | 4 Points |

|  |
| --- |
| Orienting |
| I believe I am at orienting because in regard to my individual challenge I have been thinking about the impact that any AI that I make will have on society. I have been focused on eliminating the possible negatives of my idea while also making sure that it will have a positive impact on my target users. I have discussed this with my SI teacher and have gotten a lot of feedback with regards to this topic. |
| Beginning |
| I believe I am at beginning because I have begun to to think more and more about the impact that A.I. has on society. I have also began to think about the impact that my project will have. I have expanded my general knowledge on the world of A.I. in general by reading books and listening to experts in the field. |

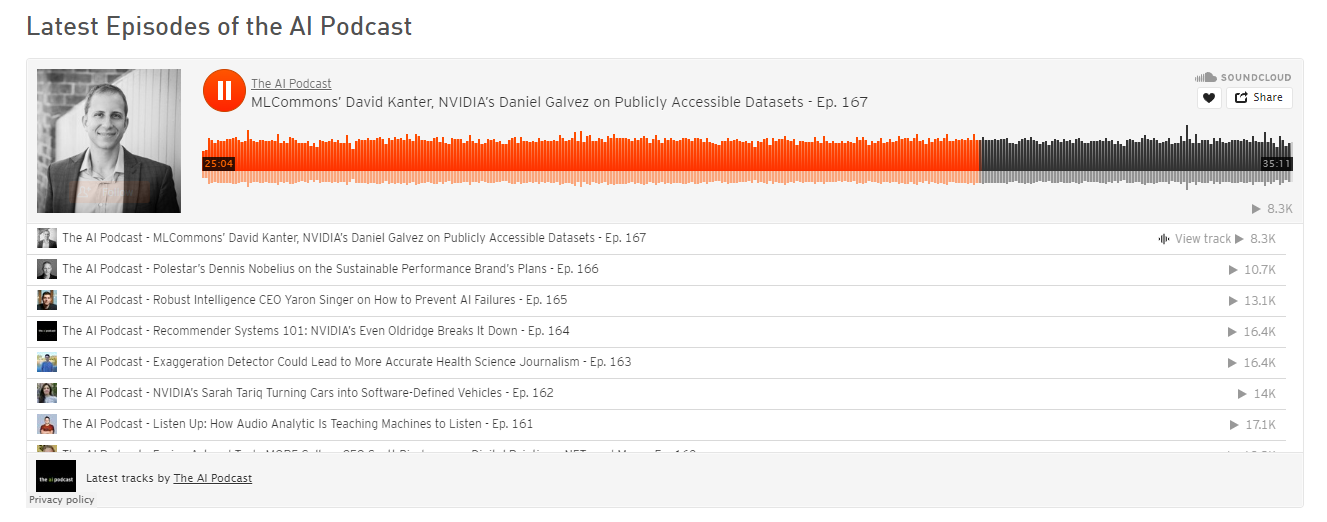
* 1. Proof



*Reflection: A book that I have read in regards to the field of A.I.*



*Reflection: I definitely think that the book helped me increase my general understanding on the field of A.I. and the different opinions on the future of this technology.*

**

*Explanation: I have also taken in interest in podcasts regarding A.I.*

## Investigative problem solving

* 1. Definition

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 analyzing your own and others’ line of reasoning.

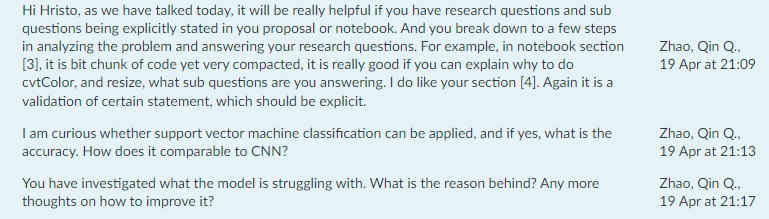
* 1. Self-evaluation

| Advanced | Proficient | Beginning | Orienting | Undefined | **Total points** |
| --- | --- | --- | --- | --- | --- |
| 4 Points | 3 Points | 2 Points | 1 Points | 0 Points | 4 Points |

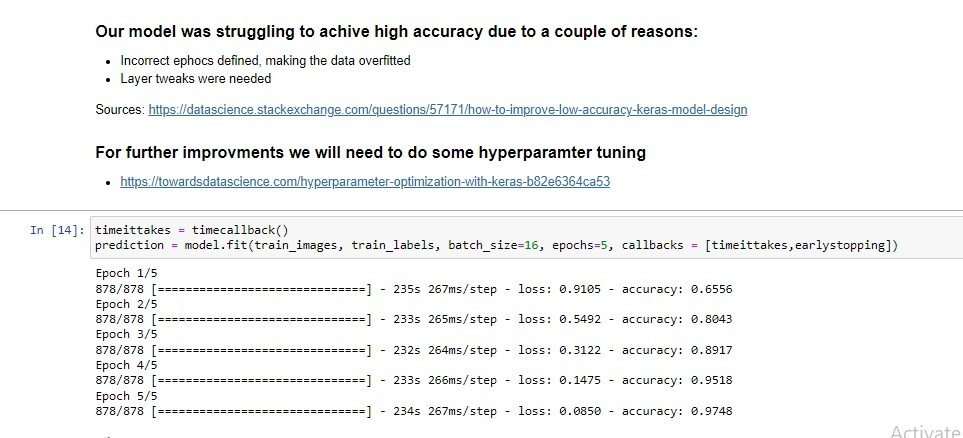
|  |
| --- |
| Orienting |
| I give myself orienting because I have done some exercises with my group project class mates in which we had to tackle a topic and ask ourselves the proper questions which we can answer doing research, and by answering them we will have a complete research document. |
| Beginning |
| I believe I am at beginning because as I have started working on my individual project and have started encountering problems with my challenge, I have become more adapt at identifying problems with my data/model and researching the appropriate methods of fixing said issues. |

* 1. Proof

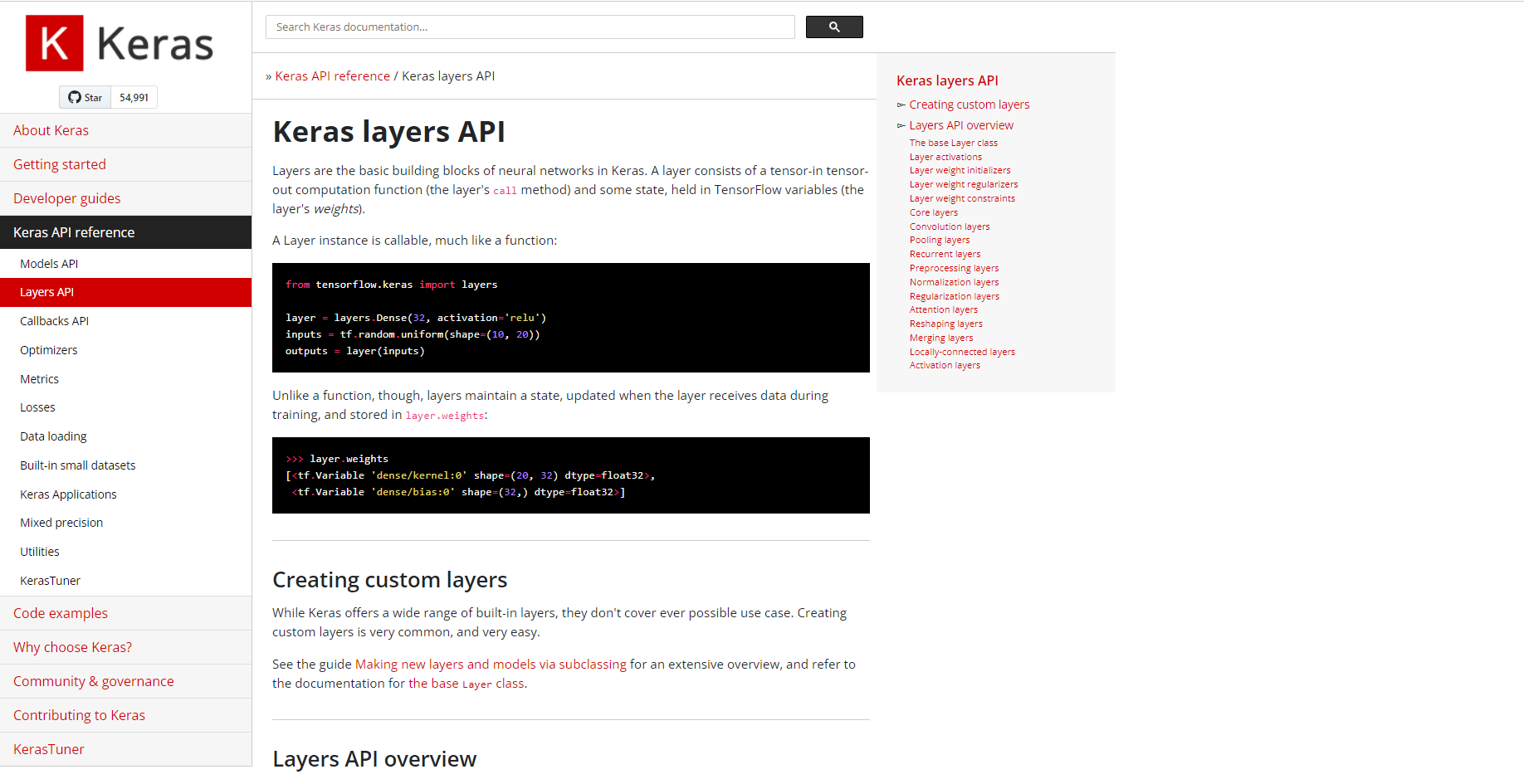




*Reflection: I definitely think that making a more clear main question in my notebook can help me explain my though and problem solving process a lot easier. I also have to make sure to explain the steps that I took to improve my model.*



*Explanation: Here I have added the reason that my model was struggling and added a source that I used to improve my model. I have also researched ways to further improve my model.*

**

*Explanation: I also looked into Keras documentation to find ways to improve my model creation to achieve better accuracy.*

## Personal leadership

* 1. Definition

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

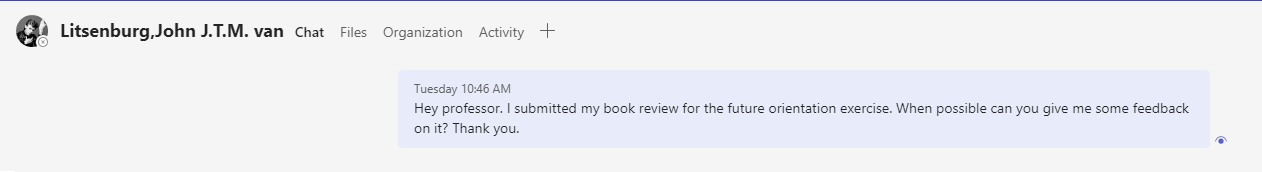
Entrepreneurial mindset includes being aware, seeing opportunities and seizing them, motivating oneself and others, being able to profile oneself, a team and others. Learning capacity means guiding your own development and study progress, showing leadership and taking responsibility, enhancing ones' 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.

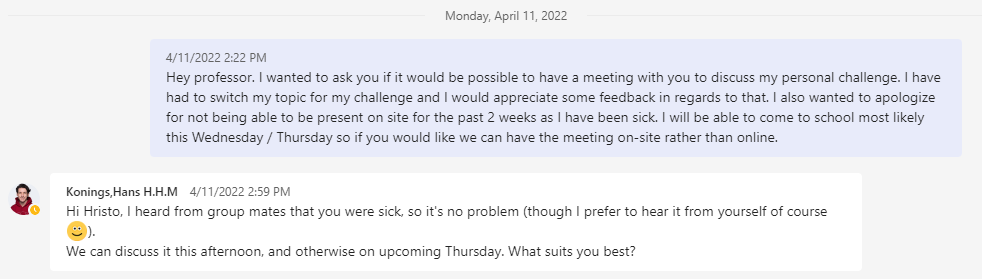
* 1. Self-evaluation

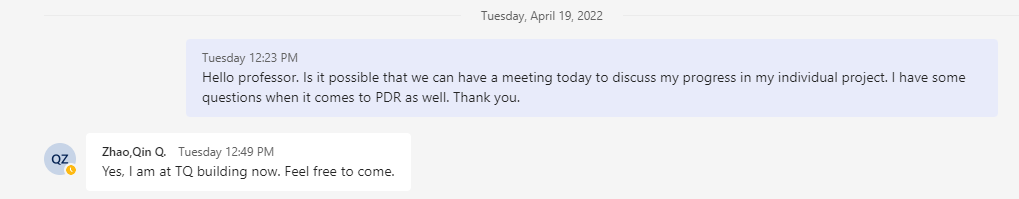
| Advanced | Proficient | Beginning | Orienting | Undefined | **Total points** |
| --- | --- | --- | --- | --- | --- |
| 4 Points | 3 Points | 2 Points | 1 Points | 0 Points | 4 Points |

|  |
| --- |
| Orienting |
| In regards to this learning outcome I believe I am still simply in the orienting phase. I come to this conclusion based not only in my development on my personal challenge and exercises but also based on the group project. While I am still in the beginning of my individual challenge and I am struggling of choosing a direction/topic out of all the possibilities, I am also orienting myself in the group project with regards to other people. Evaluating who is good at which task and how I can suggest to divide the work in such a way that everybody is doing their best. |
| Beginning |
| I believe I am at beginning because I have started to take more initiative in regards to acquiring more feedback in regards to my personal challenge. I have also been trying to study other material outside of the one provided by the school by watch videos and reading online. |

* 1. Proof

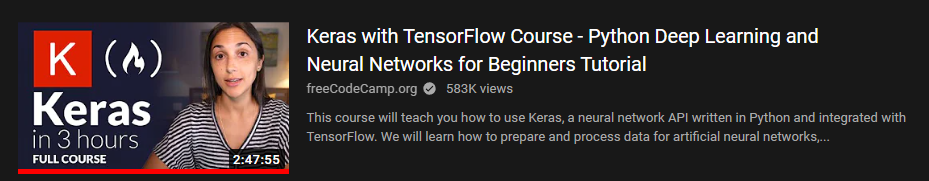






*Explanation: Taking the initiative to ask for feedback from the teachers*

|  |
| --- |
| Peer review |
| Dobri |
|  |
| Anna |
|  |
| Shanessa |
|  |
| Jonathan |
|  |
| Sharina |
|  |





*Explanation: Videos I have watched regarding 2 types of models I implement in my individual challenge.*

## OE4 Internship Preparation

* 1. Definition

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.

* 1. Self-evaluation

| Advanced | Proficient | Beginning | Orienting | Undefined | **Total points** |
| --- | --- | --- | --- | --- | --- |
| 4 Points | 3 Points | 2 Points | 1 Points | 0 Points | 4 Points |

|  |
| --- |
| Orienting |
| In regards to this learning outcome I believe I am still not fully taking the initiative to contact different companies for internship position. |

* 1. Proof

|  |  |
| --- | --- |
| Reason | Date |
| Attended Career day at fonys and talked with 5 different companies to learn more about companies | 13.04.2022 |