Project Plan

1. Five-point test

1.1. Project overview:

Maastricht University asked us to process the data they collected from adolescents who took the Five Dots Test (high school students). They already gave us a main question with a couple of sub-questions:

Main question:
- Do students use specific or diverse strategies?
- Rephrased to: **What strategies do students employ, and are they typically specific or varied?**

Sub-questions:
- *Do people with a high total score show different strategies than people with a low total score?*
- *Rephrased to:* **In individuals with high and low total scores, do we observe variations in the strategies they use?**
- *Can a combination of different aspects of the way a student makes these figures form a better predictor of the cito score or personality traits of this person?*
- *Rephrased to:* **Can a composite of various aspects related to a student's approach to these tasks serve as a more accurate predictor of their Cito score or personality traits?**

1.2. Project scope:

- Scope statement: This project is commissioned by Maastricht University to process and analyze the data collected from adolescents who participated in the Five Dots Test. The primary objective is to address the main question and associated sub-questions. The project will focus on utilizing AI and Machine Learning techniques for data analysis. The scope of this project includes:

  🟧 Phase 1 Proposal:

- Domain Understanding: Gaining a clear understanding of the domain and the research questions.
- Data Sourcing: Identifying and sourcing the necessary data for analysis.
- Analytic Approach: Defining the approach and methodology for the analysis.

🟦 Phase 2 Provisioning:

- Data Requirements: Specifying the data requirements for the project.
- Data Collection: Collecting the required data for analysis.
- Data Understanding: Gaining insights into the data through exploratory analysis.
- Data Preparation: Preprocessing and cleaning the data to make it suitable for AI and Machine Learning analysis.

🟥 Phase 3 Predictions:

- Preprocessing: Preparing the data for model development.
- Modelling: Developing AI and Machine Learning models to answer the research questions.
- Evaluation: Assessing the models' performance and their ability to answer the main question and sub-questions.

🟩 Phase 4 Delivery:

- Demonstration: Presenting the results and insights obtained from the analysis.
- Feedback: Gathering feedback and insights from stakeholders and Maastricht University.

The project will **not** include activities beyond the outlined phases and will deliver a comprehensive report summarizing the findings and model performance.

1.3. Project objectives:

Stakeholder Identification:

- Maastricht University: Those we are doing the project for, they are the main stakeholder and will be affected the most,
- Teachers: They will end up **grading** our project and its creation.
- Project group: Good cooperation and communication among themselves is necessary to achieve a successful project.

1.4. Project timeline:

The project will consist of 3 iterations

The first iteration needs to be delivered by the 29th of October.

From there we will deliver 2 more iterations, a midterm review and the final version of the project.

- Part A on the 26th of November
- Midterm review on the 3rd of December
- Part B on the 17th of December
- Final Part C on the 14th of January

1.5. Work breakdown structure: *(as of 03.11.23)*

Commented [NV1]: Each task should have a description and estimated duration.

- **Phase 1 – Proposal**
  - Project Proposal
    - The Project Proposal outlines a plan for a specific project, detailing its objectives, scope and timeline
    - The project proposal is going to take 2-3 weeks to complete
  - Project Plan
    - A project plan provides a detailed roadmap for executing a project, including tasks and timelines serving as a guide for project management and implementation.
    - The project plan is going to be updated with time
  - Domain Understanding
    - Domain understanding is the initial exploratory research step in a project, involving expert interviews, user consultations, or standard references to formulate research questions and methods, and its findings serve as a critical project proposal chapter to support decision-making and stakeholder approval
    - (period not defined yet))
  - Data Sourcing
    - Data sourcing is the initial step in AI projects, as the quality and quantity of gathered data significantly influence the performance and accuracy of AI models, emphasizing the importance of meaningful data for better results
    - (period not defined yet)
  - Analytic Approach
    - In the proposal, defining an analytic approach involves setting a specific goal for the AI project, focusing on predicting a target variable by identifying suitable data features to facilitate machine learning, ultimately culminating in the project proposal's deliverable for phase 1

- **Phase 2 – Provisioning**
  - Data Requirements
    - This serves as the first step for the collection of data. Determine what data we need to achieve the goal of the project
  - Data Collection
    - The aim of this step is to acquire a representative and informative dataset that enables the machine learning module to understand patterns, make accurate predictions and to perform other tasks
  - Data Understanding
    - In order to have a more successful outcome we need to get a better understanding of the data we are using and the insights of it using analytics and how we can improve the way they work
  - Data Preparation
    - Data Cleaning
    - Handling missing data
    - (pre)Processing data
- **Phase 3 – Predictions**
  - Preprocessing
    - Feature Selection
    - Scaling
    - Encoding
  - Modelling
    - In this step we need to select and apply various modelling techniques such as nearest neighbors, linear regression, etc.
  - Evaluation
    - Cross Validation
    - Confusion Matrix
    - Loss/score function and R^2
    - Hyperparameter tuning
- **Phase 4 – Delivery**
  - Demonstration
    - Putting our model to the test in front of one or more of our stakeholders, could be domain experts, target groups, etc.
  - Feedback
    - Feedback report