**Personal Development Report**

Learning outcomes:

1 Data Preparation & Analysis

You are able to aggregate and prepare given datasets as well as other (open) datasets and use them in data analysis and identify opportunities for predictive analytics.

2 Model Engineering

You are able to use findings from data analysis to pre-process data, apply machine learning algorithms and evaluate the quality and usefulness of produced models, for a defined domain.

3 Explainable AI

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

4 Professional Standard

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.

5 Personal Leadership

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.

OE4 Internship Preparation

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 specialisation.

|  |  |  |  |
| --- | --- | --- | --- |
| Weeks | daia | Societal impact | Machine learning |
| 1 | Thanks for submitting your thoughts on this notebook. Try to relate your arguments clearly to the 3 learning outcomes (check the detailed description). Consider all (4) the elements of data provisioning (and also think about how you would improve that). I hope you now have an idea of what a basic notebook looks like and what is important to include/improve. | In een project proposal conclusie nooit een if gebruiken, maar een defenitieve ja of nee geven. | You saw correctly that the results are different because of the train/test split. Do you know how to make them stable as well? You correctly changed the features and concluded that it did not improve the results. Adding a third feature does not change the results in this case, but in general can make it better. Note that comparing based on the plots is very subjective, it's better to use objective methods. You correctly conclude that changing n\_neighbours gives different results and understand what's happening. |
| 2 |  | De structuur van de cognetive scheme was goed. De conclusie wat we geleerd hadden (dat je niet zelf moet denken "ik weet wat de benefits zijn" en alleen dit gaan researchen ,want je weet niet alles en dan mis je informatie zoals mentale benefits) was uitstekend. | You plotted the graphs for normal distribution correctly, but mention incorrectly why the residuals should be normally distributed. You performed the explainability exercise well. You compare the two models, but make incorrect conclusions. The predictions are not close to 0, but maybe the errors are. Make sure you are clear what you are measuring and why, as this is hard to do in AI! |
| 3 | Good job, with the 4 of you you have created a first version, so now you can work from this point on to improve it regularly. Although you might have more questions than answers at this stage, you have create a first structure as starting point. It helps you to identify assumptions, raise more in-dpeth questions and of course, to start listing and descirbing the data elements you need. Looking at the headers, you have addressed some of the relevant aspects, but make sure in next versions you will go into more details. E.g. For the quality criteria, you can make your first considerations more concrete (how do you capture consistency or accuracy in a data requirement?) and once you have started to explore data (sources) you might come up with an addition to this part. You have started to make an overview of data elements, and you good broader than just the highway traffic. That is good, but makes sure you can explain why rain forecast or temperature is relevant (how is it connected to you rtarget variable?). And maybe there is more, such as fog or holidays? If you don't know if a requirement is valid or necesarry (yet), than add a comment or note and check or revise this in a later stage. Everytime you explore or analyse the data, you might need to update the requirements as well. For coming weeks: try to make the data elements more specific/detailed: think about how to include date/time aspects in order to link the data elements... Weather measures/units can also be more specific. Maybe it helps to give some examples of how this data looks like and how you would use it? E.g. is temperature the daily mean, or high/low temperature? And why you need Day if you also have Date as variable? In conclusion: this exercise was meant to help to to get started, and that is what you've done. of course, it is not a complete or perfect overview of data requirements yet; you have to keep working on that for the coming weeks, but as the start is sometimes the most difficult, you now have something to move forward from. Feel free to discuss further work or progress with me! |  |  |