Personal Development Report

ADS-A

Brent Schoenmakers | 2018-2019

Summary

This will be filled in at the end of the semester.

Inhoud

Introduction

This personal development report is for me to document my experiences and growth as a data scientist. But first, I'm going to introduce myself.

My name is Brent Schoenmakers, and I live in a village called Oisterwijk. Oisterwijk is located near Tilburg, and it takes me about 40 minutes to travel to school. I chose to ICT as my study of choice, because I was always interested in how the computer really works. I was very interested in how certain programs work, and how they might work together with different programs, and how there is rarely an error presented to the user.

I'm currently in the 4th semester of Technology, meaning that I've already done a specialization route previous semester. This specialization was Game Design. But for this semester, I wanted to try something totally different. I specifically wanted to delve deeper into the world of machine learning, because I've always found it very interesting in how a machine makes predictions with the usage of algorithms. This is the reason I chose Applied Data Science as my second specialization route.

When starting this course, I had zero knowledge of Data Science. Basically, everything that I've learned this course is completely new to me.

Learning objectives in Applied Data Science

In ADS-A there are 8 different objectives which I need to convince the teachers that I've grown in over the course of this semester. These are:

- **Reporting**
  - You are able to report in a methodologically sound way about a data analysis (plan, process documentation, report of final results, etc.).
- **Machine Learning**
  - You are able to apply machine learning algorithms for classification and regression (supervised learning) to a given data set.
- **Data Driven Organization**
  - You are able to explain what a 'data driven organisation' is, are able to argue on the maturity level of such organisation and are able to translate this into a business case for the application of data science.
- **Business Requirements**
  - You are able to translate business requirements into a structured data analysis plan.
- **Cross Validation**
  - You are able to improve the quality of machine learning models using cross validation techniques and systematic searches of the model's hyper parameters.
- **Data Quality**
  - You are able to clean data sets according to theories of data quality, in such a way that the process of cleaning is repeatable and the final result is data set suitable for data analysis.
- **Data Ethics**
  - You are aware of, and are able to reflect on your own choices in terms of the fact that laws exist regarding digital data and can explain the term "data ethics".
- **Work Ethos**
  - You are an effective co-worker in project groups, and are able to guide your own study progression by asking for, interpreting and applying feedback by teachers, tutors, coaches and fellow students.

Below I will elaborate on every single topic of the ones named above. I will describe how I have grown over the course of this semesters, and how I achieved that growth.

Learning Objective	What did I learn	How did I learn
Reporting	• I learned how to create different types of graphs.	I learned how to create my own parallel coordinates and matrix plot graphs.
Machine Learning	• I learned how the Knn algorithm works • I learned what different machine-learning types there are •	In week 1 got my hands on 3 different datasets, and through the knn-learning algorithm, I could predict certain characteristics of iris flowers, wines and computer parts. In week 2 I did a little research on the different types of machine learning algorithms and when to use what type.
Data Driven Organization	• I learned what it means to be data driven • I learned how to be data driven • I learned what it means to have clean and accessible data. • I learned about the hallmarks of a data driven organization. • Together with my group I came up with my own data drive organization with its own data driven infrastructure	In the preparation of week 2 I learned about what it means to be a data driven organization. I got my information from an article written by Carl Anderson. This article went a little in-depth on how to set up a data driven organization and what requirements you need to call yourself such a company. I also learned the difficulties of setting up your own data driven organization, and what complications it brings.
Business Requirements		
Cross Validation		
Data Quality		
Data Ethics		

Work Ethos	• I learned the value of preparation work • I learned how to make a mindmap • I learned how to differentiate different types of analytics through groupwork.	In week 2 I was tasked to compile my preparation work together with my group, and make a mindmap out of it, so we can link the information together. In class for week 2, we were tasked to come up with a business case and how to help that business by replacing an analog system with a machine learning algorithm.