CV

Florian Hein

Profile

I am a mathematician and have specialized in probability theory and statistics during my studies. Especially I dealt with the basics for machine learning and statistical simulation models. In my professional career I have further deepened these special fields. In my work as an IT consultant and later as a freelance data scientist, I have gained experience in various projects in the field of Business Intelligence and Data Science. In various technical contexts I have been responsible for data warehouse development, statistical analysis and the development of simulation and prediction models. The following section gives an overview of my previous projects and the activities I have been involved in. In the last section you will find a list of my professional career and my professional skills.

Project Experiences

VOLKSWAGEN / UMI – 8/2019 BIS 3/2020

Urban Mobility International (UMI) is a fully owned subsidiary of Volkswagen and operates the purely electric car sharing service WeShare. The company started operations in Berlin in summer 2019 and is now expanding the service to more and more cities throughout Europe. The exclusive use of electric vehicles and the resulting increased maintenance times create additional challenges for the operation of the vehicle fleet. These were also reflected in the technical requirements in the data department.

As part of the data team, my tasks included designing the data warehouse, constructing data pipelines and building statistical reports in the form of dashboards or ad hoc analyses for the various business departments in the company to support them in their work. In close cooperation with the respective stakeholders, I discussed the business requirements and tailored the database design accordingly to ensure efficient analyses.

In addition to these classical data analyses, my field of activity also included the development of algorithms for operational purposes and the programming of simulation and prediction models. Using machine learning and optimization techniques, I developed models for the prediction of demand and downtime and based on these models I wrote algorithms that control maintenance operations and

the relocation of vehicles from coldspots to hotspots. The aim was to achieve the highest possible fleet availability and utilization.

Applied technologies and qualifications: Google Cloud Platform, BigQuery, Google Data Studio, Apache Airflow, Python, Dash, Data Warehousing, Evaluation of geodata, Data Engineering, Data Analysis, Data Science, Machine Learning, Optimization

GWQ SERVICEPLUS AG, DÜSSELDORF – 11/2018 BIS 2/2019

GWQ uses statistical methods to evaluate the data of insurants of health insurance companies. The aim is to develop predictive models that identify previously undiscovered diseases amongst insured persons or determine whether persons belong to risk groups that could be affected by certain diseases in the future. The aim is to improve medical care for the insured while at the same time optimizing costs for the health insurance companies.

Within the scope of this project I have supported GWQ in various projects. I consulted on the possibilities of using Machine Learning and was involved in the implementation of these projects myself. The data was obtained from a data warehouse and prepared for the application of data science methods using SQL and R.

Applied technologies and qualifications: MS SQL Server, Machine Learning Services, R, Data Engineering, Data Science

GWQ SERVICEPLUS AG, DÜSSELDORF – 12/2018 BIS 12/2018

I gave a training course of several days at the customer's site on the subject of R and Machine Learning Services on the MS SQL Server. The main focus was to show the employees the possibilities for the use of statistical methods and machine learning procedures on the SQL Server.

Applied technologies and qualifications: MS SQL Server, Machine Learning Services, R, Data Science

GIZ, BONN UND ESCHBORN – 11/2018 BIS 2/2019

As part of a digitalization project at the customer, a study was carried out to document the existing data use and interconnectivity in the company and to reveal potential for creating added value from the existing data through the use of modern BI and data science methods. In addition, problems in data governance were to be identified and recommendations for measures to improve data quality were to be made.

I conducted interviews and workshops at the client's site and then prepared the report. In this study, the customer was provided with detailed solutions on how to optimize data governance and enable interconnected data use across departments and company sites through a cloud infrastructure with data Warehouse and Data Lake. In addition, concepts for the use of BI tools and machine learning were explained, which can solve problems that arose from the interview appointments.

Applied technologies and qualifications: Data Warehousing, Data Science, Consulting

TERRANET SOFTWARE GMBH, BERLIN – 4/2018 BIS 7/2018

The customer offers a data warehouse solution with its own software for health insurance companies. In order to meet the requirements of health insurance companies with a large number of insured persons and correspondingly large amounts of data, the performance of the ETL pipelines for the data warehouse had to be optimized. For this purpose, several stored procedures had to be reworked in order to improve the data processing durations.

Applied technologies and qualifications: Data Warehousing, Data Vault, MS SQL Server

DEUTSCHE APOTHEKER- UND ÄRZTEBANK, DÜSSELDORF – 3/2018 BIS 9/2018

The customer needed a system that would allow them to easily retrieve and read important press releases. To this end, a solution was developed in which a web crawler retrieves news reports from various sources, cleans them up and makes them available in a dashboard in a clear form for the user. In order to give the user the possibility to filter the messages according to topic or importance of the message, an artificial intelligence was integrated, which classifies all articles that have been picked up before they are made available in the dashboard.

My task in this project was to implement this artificial intelligence. During the implementation Natural Language Processing and an Ensemble method were used to solve the classification problem.

Applied technologies and qualifications: Apache Spark, Tensorflow, Machine Learning, Natural Language Processing

STADTREINIGUNG HAMBURG – 3/2018 BIS 5/2018

The data warehouse of the Hamburg city cleaning service was currently undergoing reconstruction. Step by step, the database schema was to be unified and expanded. For this purpose, ETL pipelines had to be reworked or newly created, the connection to the OLAP Cube had to be made and reports had to be generated in Excel.

I have been supporting this project for about 3 months and have developed new connections for individual departments. I have also carried out the necessary connections in the Cube as well as the generation of reports.

Applied technologies and qualifications: MS SQL Server, SSIS, SSAS, MS Excel

CONTIPARK PARKGARAGEN GMBH, DÜSSELDORF – 11/2017 BIS 12/2017

In this project, a reporting system was developed as part of a proof of concept to monitor the utilization of parking facilities. Raw data on entrances and exits were imported and processed on an SQL server. To evaluate the data, a dashboard was then set up in Power BI, which can be used to call up key figures and filter the data according to various criteria. I worked on this project completely on my own responsibility and completed it successfully.

Applied technologies and qualifications: MS SQL Server, SSIS, Power BI, Data Analysis

FASHIONETTE GMBH, DÜSSELDORF – 7/2017 BIS 12/2017

The customer is an online retailer that needed to build a data warehouse from scratch as part of the project. This was primarily designed for purchase controlling and should make it possible to determine the need for ordering goods. For this purpose a DWH was created on a MS SQL Server and a Tabular Cube was set up with Analysis Services.

The cube was used to evaluate sales numbers from the past and a calculation logic was used to compute forecasts for future requirements.

I myself was the team leader of this project and was responsible for the design of the DWH and the setup of the ETL pipelines. I also designed and implemented the calculation logic for the Cube.

Applied technologies and qualifications: MS SQL Server, SSIS, SSAS, Data Warehousing, Predictive Analysis

Professional Experiences

- Freelance data scientist and BI developer - 08/2019 till today

- adesso AG, Berlin, Consultant for Data Science & BI - 02/2018 bis 02/2019

- ixto GmbH, Berlin, Consultant for Data Science & BI - 07/2017 bis 01/2018

University Education

- University Rostock – Mathematics Master 2014 bis 2017

- University Rostock – Mathematics Bachelor 2009 bis 2014

Skills

- Programming: Python, R, Ruby, Apache Spark, Matlab, SQL, Git, HTML, CSS

- Work with machine learning frameworks e.g. Scikit-Learn, Tensorflow and Keras.

- Evaluation of geodata e.g. with h3-Framework from Uber

- Data Science: Use of various unsupervised and supervised learning methods e.g. deep learning, support vector machines, time series analyses, association analyses

- Data Engineering / ETL: Apache Airflow, SSIS

- Business Intelligence: Google BigQuery, MS SQL Server, MySQL, PostgreSQL, Power BI, Looker, Dash