

# Lebenslauf

Florian Hein

## Profil

Ich bin Mathematiker und habe mich in meinem Studium auf die Bereiche Wahrscheinlichkeitstheorie und Statistik spezialisiert. Besonders habe ich mich mit den Grundlagen für Machine Learning und statistischen Simulationsmodellen auseinandergesetzt. In meiner beruflichen Laufbahn habe ich diese Spezialgebiete weiter vertieft. In meiner Tätigkeit als IT Consultant und später als freiberuflicher Data Scientist habe ich in diversen Projekten Erfahrungen im Bereich Business Intelligence und Data Science gesammelt. Dabei habe ich in unterschiedlichsten fachlichen Kontexten Data Warehouse Entwicklung, statistische Auswertung und den Aufbau von Simulations- und Vorhersagemodellen betreut. Der folgende Abschnitt gibt einen Überblick zu meinen bisherigen Projekten und den Tätigkeiten, die ich dort übernommen habe. Im letzten Abschnitt findet sich eine Auflistung beruflicher Laufbahn und meiner fachlichen Kenntnisse.

## Projekterfahrungen

### **VOLKSWAGEN / UMI - 8/2019 BIS 3/2020**

Das Unternehmen Urban Mobility International (UMI) ist eine hundertprozentige Tochterfirma von Volkswagen und betreibt den rein elektrischen Car Sharing Service WeShare. Die Firma nahm den Betrieb im Sommer 2019 in Berlin auf und weitete den Dienst nun auf immer mehr Städte in ganz Europa aus. Durch den ausschließlichen Einsatz von Elektrofahrzeugen und die dadurch erhöhten Wartungszeiten entstehen zusätzliche Herausforderungen beim Betrieb der Fahrzeugflotte. Diese spiegelten sich auch in den fachlichen Anforderungen in der Data Abteilung wider.

Als Teil des Data Teams gehörte es zu meinen Aufgaben das Data Warehouse Design zu entwerfen, Datenpipelines zu konstruieren und für die verschiedenen Fachabteilungen im Unternehmen statistische Auswertungen in Form von Dashboards oder Ad Hoc Analysen zu bauen, um diese in ihrer Arbeit zu unterstützen. Dabei habe ich in enger Zusammenarbeit mit den jeweiligen Stakeholdern die fachlichen Anforderungen besprochen und das Datenbankdesign entsprechend zugeschnitten, um effiziente Auswertungen zu gewährleisten.

Neben diesen klassischen Datenauswertungen gehörte zu meinem Aufgabenfeld auch die Entwicklung von Algorithmen für operative Zwecke und die Programmierung von Simulations- und Vorhersagemodellen. So wurden unter Einsatz von Machine Learning

und Optimierungsverfahren Modelle für die Vorhersage von Nachfrage und Ausfallzeiten entwickelt und darauf aufbauend Algorithmen geschrieben, die Wartungsoperationen und die Relokalisierungen der Fahrzeuge von Coldspots zu Hotspots steuern. So sollte eine möglichst hohe Flottenverfügbarkeit und Auslastung erzielt werden.

**Eingesetzte Technologien und Qualifikationen:** Google Cloud Platform, BigQuery, Google Data Studio, Apache Airflow, Python, Dash, Data Warehousing, Auswertung von Geo Daten, Data Engineering, Data Analysis, Data Science, Machine Learning, Optimierung

#### **GWQ SERVICEPLUS AG, DÜSSELDORF - 11/2018 BIS 2/2019**

Bei der GWQ werden Versichertendaten von Krankenkassen mit Hilfe statistischer Methoden ausgewertet. Dadurch sollen Vorhersagemodelle entwickelt werden, die bei Versicherten bisher unentdeckte Krankheiten erkennen oder feststellen, ob Personen Risikogruppen angehören, die zukünftig von bestimmten Krankheiten betroffen sein könnten. Das Ziel dabei ist eine bessere medizinische Betreuung der versicherten bei gleichzeitiger Kostenoptimierung für die Krankenkassen.

Ich habe im Rahmen dieses Projekts die GWQ in verschiedenen Projekten unterstützt. Dabei habe ich über Möglichkeiten zum Einsatz von Machine Learning beraten und war selbst an der Umsetzung dieser Projekte beteiligt. Die Daten wurden hier aus einem Data Warehouse bezogen und mit Hilfe von SQL und R für die Anwendung von Data Science Methoden aufbereitet.

**Eingesetzte Technologien und Qualifikationen:** MS SQL Server, Machine Learning Services, R, Data Engineering, Data Science

#### **GWQ SERVICEPLUS AG, DÜSSELDORF - 12/2018 BIS 12/2018**

Ich habe bei dem Kunden eine mehrtägige Schulung zum Thema R und Machine Learning Services auf dem MS SQL Server durchgeführt. Der Schwerpunkt lag darauf, den Mitarbeitern die Möglichkeiten für den Einsatz statistischer Methoden und Machine Learning Verfahren auf dem SQL Server näher zu bringen.

**Eingesetzte Technologien und Qualifikationen:** MS SQL Server, Machine Learning Services, R, Data Science

#### **GIZ, BONN UND ESCHBORN - 11/2018 BIS 2/2019**

Im Rahmen eines Digitalisierungsprojekts beim Kunden, wurde eine Studie durchgeführt, um die bestehende Datennutzung und Vernetzung im Unternehmen zu dokumentieren und Potentiale aufzudecken, wie durch den Einsatz moderner BI und Data Science Methoden Mehrwerte aus den vorhandenen Daten geschaffen werden können. Außerdem sollten Probleme bei der Data Governance identifiziert werden und Maßnahmen zur Verbesserung der Datenqualität vorgeschlagen werden.

Ich habe selbst beim Kunden Interviews und Workshops durchgeführt und anschließend die Studie erstellt. In dieser Studie wurden dem Kunden ausführliche Lösungsansätze geliefert, wie er durch eine Cloud-Infrastruktur mit data Warehouse und Data Lake die Data Governance optimieren und die vernetzte Datennutzung über Abteilungen und Unternehmensstandorte hinweg ermöglichen kann. Außerdem wurden Konzepte für den Einsatz von BI Tools und maschinellem Lernen erläutert, durch die sich Problemstellungen lösen lassen, die sich aus den Interview Terminen ergeben hatten.

**Eingesetzte Technologien und Qualifikationen:** Data Warehousing, Data Science, Beratung

#### **TERRANET SOFTWARE GMBH, BERLIN - 4/2018 BIS 7/2018**

Der Kunde bietet eine Data Warehouse Lösung mit einer eigenen Software für Krankenkassen an. Um den Anforderungen bei Krankenkassen, mit sehr vielen Versicherten und dementsprechend großen Mengen an Daten, gerecht zu werden, musste die Performance der ETL Ladestrecken für das Data Warehouse optimiert werden. Hierfür mussten etliche Stored Procedures umgearbeitet werden, um die Abläufe bei der Datenverarbeitung zu verbessern.

**Eingesetzte Technologien und Qualifikationen:** Data Warehousing, Data Vault, MS SQL Server

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

Der Kunde benötigte ein System, über das er wichtige Pressemeldungen in einfacher Art und Weise abrufen und lesen kann. Zu diesem Zweck wurde eine Lösung entwickelt, bei der ein Webcrawler Nachrichtenmeldungen aus verschiedenen Quellen abgreift, diese bereinigt und in einem Dashboard in übersichtlicher Form für den Nutzer bereitstellt. Um dem Nutzer die Möglichkeit zu geben, die Meldungen nach Thema oder Wichtigkeit der Meldung zu filtern, wurde eine künstliche Intelligenz integriert, die alle abgegriffenen Artikel klassifiziert, bevor sie im Dashboard bereitgestellt werden.

Meine Aufgabe in diesem Projekt bestand in der Implementierung der künstlichen Intelligenz. Bei der Umsetzung wurde Natural Language Processing und eine Ensemble Methode für die Lösung des Klassifizierungsproblems eingesetzt.

**Eingesetzte Technologien und Qualifikationen:** Apache Spark, Tensorflow, Machine Learning

### **STADTREINIGUNG HAMBURG - 3/2018 BIS 5/2018**

Bei der Stadtreinigung in Hamburg befand sich zur angegebenen Zeit gerade das Data Warehouse im Umbau. Schritt für Schritt sollte das Datenbankschema vereinheitlicht und erweitert werden. Hierfür mussten ETL-Strecken umgebaut oder neu erstellt werden, die Anbindung in den OLAP Cube vorgenommen werden und Berichte in Excel generiert werden.

Ich habe bei diesem Projekt etwa 3 Monate unterstützend mitgearbeitet und habe im Rahmen dieser Tätigkeit Anbindungen für einzelne Fachbereiche neu entwickelt und sowohl die nötigen Anbindungen im Cube als auch die Erstellung der Berichte vorgenommen.

**Eingesetzte Technologien und Qualifikationen:** MS SQL Server, SSIS, SSAS, MS Excel

### **CONTIPARK PARKGARAGEN GMBH, DÜSSELDORF - 11/2017 BIS 12/2017**

Bei diesem Projekt wurde im Rahmen eines Proof of Concept ein Reporting System entwickelt, über das sich die Auslastung von Parkhäusern überwachen lässt. Rohdaten über Ein- und Ausfahrten wurden auf einen SQL Server importiert und aufbereitet. Zur Auswertung der Daten wurde dann ein Dashboard in Power BI aufgesetzt, über das

sich Kennzahlen abrufen lassen und die Daten nach diversen Kriterien gefiltert werden können. Dieses Projekt habe ich komplett eigenverantwortlich bearbeitet und erfolgreich abgeschlossen.

**Eingesetzte Technologien und Qualifikationen:** MS SQL Server, SSIS, Power BI, Data Analysis

#### **FASHIONETTE GMBH, DÜSSELDORF - 7/2017 BIS 12/2017**

Bei dem Kunden handelt es sich um einen Onlinehandel, für den im Rahmen des Projekts ein Data Warehouse von Grund auf aufgebaut werden sollte. Dieses wurde vorwiegend für das Einkaufscontrolling konzipiert und sollte es ermöglichen den Bedarf für Warenbestellungen zu ermitteln. Zu diesem Zweck wurde ein DWH auf einem MS SQL Server erstellt und mit Analysis Services ein Tabular Cube aufgesetzt. Über den Cube wurden Verkaufszahlen aus der Vergangenheit ausgewertet und über eine Berechnungslogik Prognosen für die zukünftigen Bedarfe errechnet.

Ich selbst hatte bei diesem Projekt die Teamleitung und habe das Design des DWH und den Aufbau der ETL Strecken verantwortet. Außerdem habe ich die Berechnungslogik für den Cube konzipiert und umgesetzt.

**Eingesetzte Technologien und Qualifikationen:** MS SQL Server, SSIS, SSAS, Data Warehousing, Predictive Analysis

## Berufserfahrungen

- Freiberuflicher Data Scientist und BI Entwickler - 08/2019 bis heute
- adesso AG, Berlin, Consultant für Data Science u. BI - 02/2018 bis 02/2019
- ixto GmbH, Berlin, Consultant für Data Science u. BI - 07/2017 bis 01/2018

## Hochschulbildung

- Universität Rostock - Mathematik Masterstudium 2014 bis 2017
- Universität Rostock - Mathematik Bachelorstudium 2009 bis 2014

## Kenntnisse

- Programmierung: Python, R, Ruby, Apache Spark, Matlab, SQL, Git, HTML, CSS
- Arbeit mit Machine Learning Frameworks z.B. Scikit-Learn, Tensorflow und Keras.
- Auswertung Geodaten z.B. mit h3-Framework von Uber
- Data Science: Einsatz diverser Unsupervised und Supervised Learning Verfahren z.B. Deep Learning, Support Vector Machines, Zeitreihenanalysen, Assoziationsanalysen
- Data Engineering / ETL: Apache Airflow, SSIS
- Business Intelligence: Google BigQuery, MS SQL Server, MySQL, PostgreSQL, Power BI, Looker, Dash