



Mohammed Aminul Haque

Data Engineering(MSc)

Johanna-Stegen-Straße 8

12167 Berlin, Germany

+49 – 015770964898

md.aminul.haque.7@gmail.com

Websites: <https://hawk06955.wordpress.com/> & <https://www.mdaminulhaque.com> (AI/LLM website)

LinkedIn: <https://www.linkedin.com/in/mohammed-aminul-haque-335a124b/>

SUMMARY

Experience: 9+ years of experience in IT/Software industry as a Senior Data Engineer

Education: BTech Computer Science & Engineering, MSc in Data Engineering

Interests: Data Engineering, AI/ML and Agents, Analysing & Solving any Data Problem

WORK HISTORY

Bauer Xcel Media

Jan 2023 – Present
Berlin, Germany

Senior Data Engineer – Data & Analytics (Full-time job)

- First Senior Data Engineer in the German division of Bauer Xcel Media (part of Bauer Media Group), supporting 12 Popular German digital publishing brands.
- Transformed a fragmented data landscape into a scalable GCP-based platform, reducing complexity, cost and data volume by 60%+.
- Built the Marketing Scorecard Project, automating reporting across all brands and optimizing €12M+ marketing spend and managing 32 Billion paid impressions with 800+million paid pageviews over 3+ years.
- Built unified, scalable data models integrating GA4, GSC, JW Player, CMS, and affiliate data across 12 German brands into a single solution, enabling insights across marketing, SEO, content, traffic, and revenue (2B+ yearly pageviews).
- Delivered real-time GA4 reporting (DE & UK) and a global revenue targets framework for standardized performance tracking.
- Engineered a 700k+ DE & UK Content AI assets dataset by unifying, embedding and vectorizing Contents across 12 Brands, 3 CMS systems, 2 countries, and multiple Content asset types, enabling advanced analytics and AI use cases.
- Developed a multilingual(50+ languages) RAG-based AI Tool (Multilingual Semantic Vector search + OpenAI + Hugging Face Embedding Models + similarity logic) supporting 9+ very sophisticated Content related AI business use cases solved via Vectors + OpenAI API and hosted tool on Kubernetes.
- Pioneered modern workflows using GCP, GitHub, Terraform, Airflow DAG's (Astronomer + Cloud Composer), Pub/Sub, Cloud Functions, Cloud Run etc
- Delivered Projects for Ströer advertising & Nielsen AGF Traffic data.
- Building multiple prototypes: OCR based Trend recommendation, Web Parsing, Brand based Recommendations of Editorial Actions to do and New Content Topic Ideas to write using AI Agents, OpenAI and vectors, Extracting Trends of internal website Traffic via Wordcloud etc

X24Factory GmbH

Sep 2021 – Dec-2022
Berlin, Germany

Data Engineer – BI & Marketing (Full-time job)

- Owning the complete Data Engineering technology stack & working with GCP, Python, BigQuery as DWH, Cloud Storage Buckets, SQL, Cloud functions, Cloud Scheduler, IAM, Kubernetes, BigQuery extracts of Google Ads and Google Analytics data, Gitlab, DBT Cloud, CSCart, PostgreSQL, Ingenious, Supermetrics.
- Being the only Data Engineer working closely with C levels, Marketing Performance Managers, BI Analysts, Backend, Products, Sales and Key Account management team to understand the requirements and build data pipelines to make high quality data available in DWH for easy, fast analysis by everyone.
- Understanding of & exposure to clicks and cpc/cpo based ecommerce business

models for retail media business and conversion based marketplace business and tools/services like Google Merchant Center, Channel Pilot.

- Extensively worked with Google Analytics and Ingenious tracking data and also worked on automating and adjusting Google Ads campaigns and Product Groups through Python generated Products outbound feed files.

Air Up GmbH

Feb 2021 – Jul 2021
Berlin, Germany

Data Engineer / BI Engineer (Full-time job)

- Joined as a BI Engineer and owning the Data Engineering technology stack including PostgreSQL DWH hosted in Azure and building data interfaces to make data available to Business stakeholders, Data analysts, Data Scientists & Marketing Analysts for analytics, business insights or datawarehousing.
- Taking care of all data integration activities like using Fivetran as ELT tool and integrating data from diverse sources like Google analytics, Shopify, Amazon, Klaviyo, csv/txt files, API, JSON/XML etc. Using Google Cloud Platform(Cloud Functions, Cloud Repositories, Big Query), Python and SQL queries to collect raw data from non standard sources or DWH and transform them using DBT into direct consumable data format for BI reporting requirements.

Jacobs University

Jan 2020 – Aug 2020
Bremen, Germany
marketing

Data Scientist/Machine Learning Engineer (Student Assistant / Part-time job)

- Used Statistical data analysis, data preprocessing and Predictive modeling using Machine Learning algorithms for finding factors affecting Cocoa prices.
- Developed Statistical, Machine learning & Deep learning Forecasting models

Cisco

April 2017 - Aug 2017
Bangalore, India

Senior Software Engineer (Full-time job)

- Worked in a Team lead role and supported Cisco's automation product named Tidal which is used by Companies for their business process automation.
- Solved Cisco's Tidal related tickets and incidents of Customers. Also raised software bugs and worked with the Software Development team to resolve.

Accenture

Dec 2013 – Jan 2017
Bangalore, India

Application Development Analyst (Full-time job)

- Worked as a developer in Sales & Marketing Business Intelligence (BI) and Data integration team and coded and implemented the end to end BI systems, Data Warehouse (DWH) and Data models & pipelines for Global CRM rollout.
- Tuned and optimized processes or pipelines for higher business productivity.
- Identified loopholes or issues in data processes. Examples: Identified and resolved a bug in code which was causing Sales orders getting lost during the daylight saving period. Parallelized a critical data pipeline generating 96 reports for 6 countries by tuning ETL processes, which resulted in a time reduction of 60-70% for the critical data pipeline.
- Being a critical developer resource, technically helped the Support team in case they were not able to resolve a high priority (Level 1/2 incident) quickly.
- Helped Data analytics team with ad-hoc data reports using SQL queries.
- Awarded Accenture Delivery Excellence award

EDUCATION

Jacobs University

Sep 2018 -Jul 2020
Bremen, Germany

Master of Science in Data Engineering

GPA: 1.72 , Awarded Admission Scholarship.

Courses: Data Analytics, Data Mining, Machine Learning, Big Databases & Cloud Services, Data Visualization & Image Processing, Current topics in Data Engineering, Web Analytics, Data analytics in SCM, Data management for Graduate Students, Big Data challenge, Data Acquisition Technologies & Sensor Networks, Systems Biology, Network Approaches in Biology & Medicine, Advanced Projects 1 & 2.

NIT Silchar
Sep 2009 – May 2013
Silchar, India

Bachelor of Technology in Computer Science & Engineering

Awarded Merit-cum-Means Scholarship by Ministry of Minority Affairs, Government of India

SKILLS

Programming Languages	Python, SQL, R
AI/ML and Vector Model	LangChain, Prompt Engineering, AI Agents, Vector Databases, Vector Search, LLMs (OpenAI, Claude, Gemini, Mistral, Hugging Face, local models), Embeddings, RAG
Frameworks / Libraries	R Shiny, Keras, TensorFlow, Apache Spark, NumPy, Pandas, Scikit-learn, Dplyr, Caret, Plotly, Anaconda, DT, iGraph
Cloud services	Google Cloud Platform(Cloud Run, Cloud Functions, Cloud Storage, GKE, Cloud Composer, Cloud Build, Secret manager etc) Astronomer, DBT Cloud, Google Ads, Fivetran, Github, Google Cloud Repositories, Salesforce, Veeva, ServiceNow, HP ALM
Tools	Docker, GitHub, Kubernetes, Terraform, Airflow, Unix Command Line, PVCS Version Manager, Postman, Workbench for Salesforce API, Apex Data Loader, Java VisualVM, fastThread TDA (Thread Dump Analyzer)
IDE	RStudio, Jupyter, PyCharm, Spyder
BI/Reporting Tools	Tableau , Klipfolio, Google Data Studio, Power BI
Software/Application	WinSCP, Putty, Informatica PowerCenter, Informatica Cloud Services, Tidal, AutoSys, Microsoft Excel & Word
Databases/Storage/DBMS	BigQuery, MySQL, Oracle, Postgres, NoSQL, MongoDB, Microsoft SQL Server, MySQL, Snowflake, Workbench, Toad for Oracle, REST API, documents (csv, tsv, text)
Spoken Languages	English: Fluent (CEFR C1) German: Basic (CEFR A2-B1)

PROJECTS/TRAINING

Training, 2020	24 Badges and 7 Trophies completed in Microsoft Learning website for Data Analyst role track and Azure Data Engineering track. Link to my Microsoft Learning badges : https://docs.microsoft.com/en-us/users/mohammedaminulhaque-1375/achievements
Hobby, 2020	Developed a free cloud-based Data Visualization tool for everyone to use : https://hawk06955.shinyapps.io/MDs_Visualization_tool/
Hobby, 2020	Developed a cloud-based Business Influence Network analysis tool to graphically understand the Business Network of Customers and Sellers by combining business data with a scientific network analysis approach : https://hawk06955.shinyapps.io/business_influence_network/
Hobby, 2020	Developed a scientific Human Disease Network Visualization tool : https://hawk06955.shinyapps.io/human_disease_network/
Academic, 2020	Master Thesis: Identifying factors influencing Cocoa prices using Predictive Modelling and a comprehensive study of Statistical, Machine Learning and Deep Learning Forecasting models to forecast the daily Cocoa returns
Academic, 2019	Advanced Project 2: Using Statistical analysis, Machine Learning & Deep Learning Classification and Forecasting models in Financial time series data
Academic, 2019	Using Artificial Intelligence in Supply Chain Management for Market and Demand segmentation of Used cars in the US market
Academic, 2019	Integrating IoT in Logistics: Temperature sensor monitoring for Shipments
Academic, 2019	Implementing machine learning algorithms for digits classification using MNIST data
Academic, 2019	Using Classification based Machine Learning algorithms for Human Activity Recognition (HAR) using 19 electronic sensors

- Academic, 2019 Advanced Project 1: Developed a cloud-based BI tool named AutoBI for intelligent automation and self-service business data analysis and reporting :
<https://hawk06955.wordpress.com/autobi-tool/>
- Academic, 2018 Using log analyzing tools (Nihuo Web Log Analyzer & Sawmill) and a tag-based tool (StatCounter) for web analytics
- Professional, 2013-2017 Multiple projects designed, developed, tested, deployed, migrated, automated/scheduled, UAT, Go-Live and later supported the BI and data processes/pipelines/interfaces for Global CRM rollout. Examples of few projects are building the interfaces of users, customer master data, products, product metrics, alignment data, multichannel cycle plans for marketing, data consumption & delivery automation of files/DB/Salesforce/REST API & SOAP Web Services(XML/JSON), Shell scripting and using Job Schedulers for business process delivery and monitoring.