



Rana Hamza Intisar

📍 Berlin, Germany

☎ +49 15143622712

✉ ranahamzaintisar@gmail.com

🔗 Github Repository

🌐 LinkedIn

🌐 Personal Website

Nationality: Pakistani

WORK EXPERIENCE

May 2021 – Present

Senior Data Engineer

LexisNexis IPlytics GmbH, Berlin, Germany

- Migrated legacy C code to Databricks using PySpark and Azure DevOps, enhancing batch data processing efficiency
- Designed, implemented and optimized scalable, high-performance data pipelines to accommodate increasing data volumes
- Developed and maintained Spark-based data pipelines powering the live platform (IPlytics)
- Designed and contributed to the migration of legacy systems to a modern, state-of-the-art architecture, improving scalability and performance
- Orchestrated and scheduled Spark jobs on Kubernetes using Apache Airflow, ensuring efficient workflow execution
- Optimized Spark jobs to enhance performance and mitigate bottlenecks, improving overall data processing efficiency
- Led the migration of a key data source from PostgreSQL to Delta Lake, enhancing data reliability and scalability
- Designed, built, and maintained data processing pipelines for the Data Lake using Delta Lake on Azure Cloud
- Developed NLP-based algorithms utilizing Gensim to compute similarity matrices for text comparison, and derived similarity scores based on the computed matrices to enhance text analysis and information retrieval
- Optimized Agile workflows to enhance productivity, collaboration, and timely delivery while driving continuous improvement and feature development
- Played a key role in the hiring process for Data Engineers and Data Scientists, overseeing onboarding, mentoring, and skill development within the team

February 2021 – April 2021

Data Scientist - Artificial Intelligence

Afiniti, Islamabad, Pakistan

- Developed generative models to optimize customer-agent pairing in call centers
- Designed and implemented data pipelines for real-time ML model deployment

July 2019 – June 2020

Data and Machine Learning Engineer

Alvary (A company of Allianz), Hannover, Germany

- Developed a demo use case for the Granary platform using the Opensky Network live API, implementing micro-services for real-time data processing
- Deployed containerized applications using Docker and Kubernetes on cloud platforms
- Built data pipelines with Spring Boot for efficient ETL processing
- Converted real-time data into time series and stored it in InfluxDB
- Implemented unsupervised learning techniques for anomaly detection in live data streams
- Created data visualizations using Metabase and Grafana to support decision-making
- Contributed to Agile sprints, enhancing the Granary platform's development and scalability

March 2019 – May 2019

Data Analyst - Internship

Lobensberg Gute Weine, Bremen, Germany

- Developed a product recommendation system using predictive analytics and machine learning algorithms (TensorFlow, GCloud Analytics Platform) to provide customers with personalized wine recommendations
- Conducted analysis on reactivating dormant customers by offering tailored promotions and bundles based on previous purchasing trends
- Independently analyzed and summarized findings, delivering comprehensive presentations to the team

August 2017 – August 2018

Data Engineer - Professional Services Consultant

Teradata Global Delivery Centre, Islamabad, Pakistan

- Utilized Big Data infrastructure to develop Data Lake and Triggers for clients such as Jazz, Banglalink, and DJezzy, leveraging technologies like Hive, Spark, Ignite, Cassandra, Kafka, Nifi, and Tez
- Created Source-to-Target mappings to transform data according to client specifications
- Conducted regular client meetings to gather requirements and ensure alignment with development objectives
- Tested scripts, scheduled, and monitored jobs to ensure reliable execution
- Developed solutions with Teradata UDI Studio to enhance data integration
- Participated in User Acceptance Testing (UAT) to validate and deploy release packages

EDUCATION

August 2018 – June 2020

Masters of Science in Data Engineering

Jacobs University, Bremen, Germany

Focusing on Data Analytics, Data Mining, Machine Learning, Statistical Modeling, Big Databases and cloud services, Databases, Data Visualization and Image Processing, Data Security

August 2013 – July 2017

Bachelor of Science in Computer Sciences

National University of Computer and Emerging Sciences, Islamabad, Pakistan

Focusing on Data Structures, Advanced Programming, Database Systems, Artificial Intelligence, Data Warehousing, Concurrent and Distributed systems

PROJECTS

February 2020 – June 2020

Anomaly Detection Implementation

Alvary (A company of Allianz), Hannover, Germany

- Performed data ingestion, transformation, and storage using Java Spring Boot, Docker, and Kubernetes, with deployments on AWS and Azure cloud platforms
- Generated metrics for hourly and daily windows, running anomaly detection models on live data streams, and deploying Docker images on Kubernetes using Python

LANGUAGE SKILLS



TECHNICAL SKILLS

Programming Languages	Java (SE+EE+FX), Python, R, C++, GO, SQL, UNIX Shell
Databases	MySQL, Oracle, TeraData, SQLite, PostgreSQL, InfluxDB, HiveQL, Cassandra, DeltaLake, Prometheus
Big Data Tools	Hadoop, Spark, Hive, Pig, Sqoop, Kafka, MapReduce, Ignite, Airflow, Solr, Elastic Search
Data Science Tools	Pandas, NumPy, SciPy, Scikit-learn, Keras, Matplotlib, Tensorflow
Software	Databricks, Google Kubernetes, Helm, Linux, Docker, Oracle, Microsoft SQL Server, Microsoft Azure, AWS, Google Cloud, Graffana, Metabase, Tableau, Dremio

Licenses/Certifications

March 2025	Databricks Data Engineer Associate - O'Reilly
June 2024	Data Engineering on Databricks - Databricks
June 2024	Data Analysis on Databricks - Databricks
February 2023	Kubernetes Certified Application Developer - Udemy
November 2022	Apache Airflow: The Operators Guide - Udemy
January 2022	Data Analytics with PySpark - Udemy
April 2021	AWS Foundations: Getting Started with the AWS Cloud Essentials - Pluralsight
April 2021	Getting Started with Spark 2 - Pluralsight
April 2021	Programming with R - Pluralsight
May 2018	Apache Ignite and Apache Cassandra Training - Teradata