Rana Hamza Intisar

✉ hamzaintisar@hotmail.com 📞 +49 15143622712 📍 Berlin, Germany in Linkedin 🔗 Personal Website

🐙 Github Repository

💼 WORK EXPERIENCE

05/2021 – Present
Berlin, Germany

Senior Data Engineer, *LexisNexis IPlytics GmbH*

- Modernized a mission-critical legacy system by **migrating C# based batch jobs to Databricks with PySpark**, improving maintainability and **reducing processing time by ~50%**
- Designed and optimized **high-throughput, scalable ETL pipelines** handling **3–5× growth in data volume**, ensuring end-to-end reliability and low-latency delivery
- Developed and maintained **Spark-based streaming and batch pipelines** powering live platform features, improving **data freshness from daily to near real-time**
- Led the architectural shift from monolithic workloads to a **cloud-native distributed design**, improving **system availability (99.9% SLA)** and future scalability
- Built and productionized **Kubernetes-orchestrated Spark workflows with Apache Airflow**, cutting operational interventions by **40%** through automation
- Profiling and resource tuning improved Spark cluster usage efficiency, reducing cost per run by **30%** and eliminating key performance bottlenecks
- Drove the migration of a core data source from **PostgreSQL to Delta Lake**, unlocking **ACID guarantees**, improved query speeds, and zero-downtime evolution
- Designed and maintained **Delta Lake pipelines on Azure Data Lake**, improving governance, lineage tracking, and downstream developer autonomy
- Created **NLP-based semantic similarity models** (Gensim) improving text-analysis accuracy and information-retrieval relevance by **>25%**
- Accelerated delivery by introducing Agile workflow optimizations, improving sprint predictability and throughput by **~20%**
- Supported team scaling — **hiring, onboarding, and mentoring** data engineers and data scientists, raising overall delivery quality and standards within the team

02/2021 – 04/2021
Islamabad, Pakistan

Data Scientist, *Afiniti*

- **Developed and productionized generative AI models** to optimize customer–agent pairing in call centers, improving **first-call resolution by ~15%** and reducing **average handling time by 10–20%** through smarter skill-based routing
- **Designed and built real-time ML deployment pipelines** enabling seamless **model updates, low-latency inference**, and **automated monitoring**, significantly improving model uptime and operational reliability

07/2019 – 06/2020
Hannover, Germany

Data and Machine Learning Engineer, *Syncier GmbH (A company of Allianz)*

- Engineered scalable real-time data ingestion and transformation pipelines using **Java Spring Boot** and **Python**, reducing processing latency and improving data delivery reliability across multiple cloud environments (**AWS & Azure**)
- Built microservices consuming the **OpenSky Network live API**, converting aviation telemetry into time-series datasets stored in **InfluxDB**, enabling continuous operational analytics
- Designed and deployed **containerized applications** with **Docker** and **Kubernetes**, improving deployment repeatability and streamlining CI/CD workflows
- Applied **unsupervised anomaly detection** models on live data streams, generating hourly and daily performance metrics that improved early-issue detection and system visibility
- Developed monitoring dashboards and analytics visualizations in **Metabase** and **Grafana**, enabling real-time decision-making for product stakeholders
- Actively contributed to Agile ceremonies and cross-functional delivery, driving feature releases and platform scalability improvements for the Granary product

03/2019 – 05/2019
Bremen, Germany

Data Analyst - Internship, *Lobensberg Gute Weine*

- **Built and deployed a personalized product recommendation engine** using TensorFlow and Google Cloud Analytics, driving **+18–25% uplift in click-through-rates** and increasing **cross-sell conversions** through predictive customer preference modeling
- **Analyzed dormant customer behavior** and developed targeted promotion strategies based on historical purchase patterns, contributing to **reactivation of churned users** and **boosting repeat-purchase rates**

- **Independently delivered data-driven insights** through clear, narrative-focused presentations, shaping **marketing campaign decisions** and improving stakeholder alignment on customer engagement strategies

08/2017 – 08/2018
Islamabad, Pakistan

Data Engineer, *Teradata Global Delivery Centre*

- **Built and optimized large-scale Data Lake architectures** for Tier-1 telecom operators (Jazz, Banglalink, Djezzy), integrating distributed technologies including Hive, Spark, Ignite, Cassandra, Kafka, NiFi, and Tez to enable **high-throughput ingestion and analytics**
- **Engineered Source-to-Target data mappings** to ensure accurate business rule implementation and seamless downstream consumption across enterprise systems
- **Led requirements discovery sessions with client stakeholders**, translating business objectives into scalable technical designs and ensuring **delivery alignment**
- **Automated job scheduling, testing, and operational monitoring**, improving **pipeline reliability** and reducing execution failures
- **Enhanced data integration workflows** by developing robust ETL components in Teradata UDI Studio, reducing **manual processing overhead**
- **Supported full release lifecycle** including UAT, validation, and controlled deployments, ensuring **smooth production rollout** and better feature adoption

🎓 EDUCATION

08/2018 – 07/2020
Bremen, Germany

Masters of Science in Data Engineering, *Jacobs University*

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

08/2013 – 07/2017
Islamabad, Pakistan

Bachelor of Science in Computer Sciences,
National University of Computer and Emerging Sciences

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

🧠 TECHNICAL SKILLS

Programming Languages

Java, Python, R, C++, GO, SQL, UNIX Shell

Big Data Tools

Hadoop, Spark, Hive, Pig, Sqoop, Kafka, MapReduce, Ignite, Airflow, Solr, Elastic Search

Software

Databricks, Snowflake, Google Kubernetes, Helm, Linux, Docker, Oracle, Microsoft SQL Server, Microsoft Azure, AWS, Google Cloud, Graffana, Metabase, Tableau, Dremio

Databases

MySQL, Oracle, TeraData, SQLite, PostgreSQL, InfluxDB, HiveQL, Cassandra, DeltaLake, Prometheus

Data Science Tools

Pandas, NumPy, SciPy, Scikit-learn, Keras, Matplotlib, Tensorflow

📑 PROFESSIONAL COURSES

- Python Essentials - O'Reilly
- Data Engineering on Databricks - Databricks
- Kubernetes Certified Application Developer - Udemy
- Data Analytics with PySpark - Udemy
- Getting Started with Spark 2 - Pluralsight
- Apache Ignite and Apache Cassandra Training - Teradata

- Databricks Data Engineer Associate - O'Reilly
- Data Analysis on Databricks - Databricks
- Apache Airflow: The Operators Guide - Udemy
- AWS Foundations: Getting Started with the AWS Cloud Essentials - Pluralsight
- Programming with R - Pluralsight

🌐 LANGUAGE SKILLS

Urdu	English	German
Native	C1 Level	B1 Level