Mustafa Lakhani

Date of birth: 06/10/1999 **Phone:** (+49) 15753616464

Email: mustafa.ml.lakhani@gmail.com

Home: Gabain str 5, 12247 Berlin (Germany)

ABOUT MYSELF

Dedicated engineer with a passion for leveraging state-of-the-art techniques to solve real-world problems. Experienced in data engineering, cloud architecture, big data analysis, microservice architecture, automation, data analysis, statistical analysis, data mining, and predictive modeling.

WORK EXPERIENCE

Data/Dev Ops Engineer

Dryad Networks [15/11/2022 – Current]

City: Berlin | **Country:** Germany

- **Designed & maintained automated cloud-native data pipelines** using **Python**, **AWS** for collecting sample data to train advanced machine learning algorithms, including **Transformer Models**, **CNNs**, **LSTMs**, **ConvLSTMs**, **Optical Flow algorithms**, and **OpenCV** for wildfire detection.
- **Utilized microservice architecture** with **Docker** and **Kubernetes** to deploy scalable and resilient services, ensuring high availability and fault tolerance.
- **Employed Terraform** to design and deploy cloud architectures as infrastructure as code (**IaC**) on **AWS**, automating resource provisioning and management.
- **Built real-time streaming data pipelines** using **Apache Kafka**, enabling the processing and analysis of live data streams for time-sensitive insights.
- **Used Prometheus and Zookeeper** to manage and monitor the Kafka streaming pipelines, ensuring stability, fault tolerance, and efficient operation of the data streams.
- **Implemented CI/CD pipelines** with **Jenkins**, streamlining the deployment process and ensuring continuous integration and delivery for multiple data engineering projects.
- **Processed large volumes of data in batches** using **Apache Spark**, optimizing the data flow and enabling efficient handling of big data workloads.
- **Utilized dbt** (Data Build Tool) to transform and model data, enhancing the robustness and maintainability of data pipelines.
- **Established and maintained data dashboards** using **The Elastic Stack**, **SQL**, and **Python** for data cleaning and visualization, providing actionable insights for the optimization of machine learning models.
- **Created and maintained financial data pipelines** to assess and estimate cost savings, utilizing advanced data processing techniques. Graphed and presented these insights to customers, enhancing customer satisfaction through clear communication of value.
- **Developed and maintained Power BI visualizations** and other reports for the documentation team, simplifying complex data for stakeholders and making it easily understandable and actionable.
- **Applied Agile methodology** to manage projects, ensuring regular sprints, continuous delivery, and collaboration with cross-functional teams, leading to quicker iterations and more efficient project management.
- **Proactively identified process improvements** and innovation opportunities within the data analysis function, leading to a 20% increase in operational efficiency.
- **Fine-tuned and optimized pre-trained open-source LLMs** and other **ML models** to enhance model accuracy and relevance for specialized applications.
- **Collaborated with key stakeholders** to ensure that data analysis efforts were aligned with business goals and supported critical decision-making processes.

Intern for ML research

CERI Schweinfurt [03/2021 – 10/2022]

City: Schweinfurt | **Country:** Germany

- Utilized **Machine Learning** to allocate production tasks among workers and automated machines, **achieving maximum efficiency.**
- **Developed pipelines to collect data from STL files, extracting critical information** such as dimensions, weight, and tolerances, **and saving them to a SQL database.**
- Collected and labeled training data from various production line scenarios to train machine learning models.
- Utilized **AWS Glue** to clean and prepare data for analysis.
- Tweaked **open-source machine learning models** like **Random Forest Classifier** to predict whether a task is best performed by workers or automated machines, **resulting in an estimated 12% increase in productivity.**

EDUCATION AND TRAINING

Bachelor of Engineering Mechatronics

Technische Hochschule Würzburg-Schweinfurt

City: Schweinfurt | **Country:** Germany

LANGUAGE SKILLS

Mother tongue(s): English

Other language(s):

German

LISTENING B1 **READING** B1 **WRITING** B1

SPOKEN PRODUCTION B1 **SPOKEN INTERACTION** B1

Levels: A1 and A2: Basic user - B1 and B2: Independent user - C1 and C2: Proficient user

DIGITAL SKILLS

Javascript / Python / Azure and AWS / Terraform / Jenkins / MongoDB / Apache Airflow / Apache Spark / Apache Kafka / Docker / Kubernetes / PostgreSQL / Prometheus / Zookeeper / Git / Business Intelligence/ Business Analytics / DBT / Elastic Stack / Power BI / Data Modelling / Machine Learning / Linux