Elaaf Shuja

elaaf.shuja@gmail.com | linkedin.com/in/elaaf | github.com/elaaf
+49-176-70369579 | Berlin, Germany

Technical Skills

Design & development of simple, scalable & maintainable backend services, distributed systems, and data pipelines.
Languages: Python, Go, C++, Rust, JavaScript, Java **Tools:** Docker, Kubernetes, Git, Jenkins, Helm, Terraform

Work Experience

Delivery Hero - deliveryhero.com — Berlin, Germany
Senior Software Engineer - Data & Backend — July 2022 - Present
Part of the Global Recommendations Team, serving recommendations to 2.2 billion users for 12 sub-brands (FoodPanda, HungerStation, PedidosYa, etc) in 70 countries.
- Developing, monitoring, and maintaining data pipelines in Python, Java (BigQuery, Dataflow), orchestrated using Airflow.
- Developing, monitoring, and maintaining team services including Serving-APIs, API-gateway, and centralized cache in Python, Golang.
- Responsible for the development, A/B testing, and rollout of cuisine recommendation strategy, yielding a +6% uplift in CVR in the A/B test.
- Worked on a dbt-airflow based feature-store to create and store features utilized in model training pipelines.
- Developed LTR-machine learning model training and deployment pipelines for serving recommendations.
- Led initiative to reduce daily operational costs by 11% by optimizing Kubernetes node-type/region, API code, and DB resources.
- Led initiative to migrate our full services stack and data pipelines to a different GCP region reducing cost by switching to nd2 machine type and reducing intra-region p95 latency by 15% for end-users.
- Being on-call for ensuring high availability of our critical recommendation services running in 2x K8s clusters in 5 global regions each.

Keyrus - keyrus.com — Pakistan (Remote)
Senior Software Engineer - Distributed Systems — August 2021 - July 2022
Development of an open-source data integration tool and data engineering consultancy work.
- As the tech lead, I helped define the product strategy from the engineering perspective to deliver key product functionality.
- Led the design and development of the open-source data integration tool, leveraging technologies such as Airbyte, Spark, Airflow, dbt.
- Built a service in Python to dynamically generate Airflow DAGs to orchestrate ETL workflows.
- Built a service to dynamically build Airflow DAGs and Tasks for DBT models incorporating their interdependencies using the dbt manifest.
- Created a Python module to perform Incremental File Ingestion between Distributed File Systems (S3, ADLS, HDFS) leveraging Spark Structured Streaming.
- Worked as an External Consultant for MAF (Majid Al Futtaim) Group operating 15 Shopping Malls in the Middle East being responsible for the data ingress/egress jobs into their Data Lake (HIVE).

ADDO AI - addo.ai — Lahore, Pakistan
Software Engineer - Big Data & Machine Learning — March 2019 - August 2021
Peta-Byte scale Data Engineering and Machine Learning projects in Telecom & Healthcare.
- Analysis of existing Teradata SQL and their conversion to PySpark jobs using Spark SQL.
- Development, unit/system-integration/user-acceptance testing of ETL pipelines (Spark jobs) of over 35 distinct business streams and 12 dimensions writing to HIVE data lake.
- Designed and Implemented the strategy for the PII data masking in the HIVE data lake.
- Optimized Spark jobs (broadcast joins, persist) transforming 700GB of data daily, reducing overall ETL run time from 5 hrs to under 3 hrs.
- Implemented data masking on PII (Personally Identifiable Information) data across the Data Platform.
- Performed User Acceptance Testing and System Integration Testing of Spark ETL jobs, verifying KPIs against the legacy system.
- Introduced KPI-based Data Quality checks and performed RCA for mismatched KPIs and introduced appropriate fixes.
- Built a Data Analytics platform on Microsoft Azure for Mercy Health a major US Healthcare provider using Databricks, and Informatica.
- Led the development effort for a Market Competition Analysis use case aimed at measuring self-growth against competitors.

Information Technology University — Lahore, Pakistan
Machine Learning Researcher — August 2018 - March 2019
Research work in Natural Language Processing (NLP) and Computer Vision at ITU in collaboration with QMUL and MIT.
- Analysis of Textual/Image misinformation propagation in closed social networks e.g. WhatsApp. ASONAM 2020 - PDF
- Unsupervised image-image translation using attention-guided cycle GANs. GitHub
- Facial Emotion Recognition using CNNs and Haar-cascade filters.

Education

M.S Computer Science | Information Technology University — 2018 - 2020

B.E Electrical Engineering | National University of Sciences and Technology — 2013 - 2017