Abdullah Ashfaq

Python, MySQL, C/C++, R, JavaScript

San Diego, CA | (619) 798-8102 | aashfaq@ucsd.edu | linkedin.com/in/abdullahashfaq11 | github.com/AbdullahAshfaq

EDUCATION

University of California San Diego — San Diego, CA
Master of Science in Data Science — Expected**: June 2025**

National University of Sciences and Technology (NUST) — Islamabad, PK
Bachelor of Electrical Engineering – CGPA: 3.99/4 — **June 2020**
- Awards: Ranked 2nd out of 180 students. Merit Scholarship for all semesters (Given to top 3 in Class)

Utah State University — Logan, UT
US Dept. of State Global UGRAD Semester Exchange Scholarship — **May 2018**

SKILLS

Languages: Python, MySQL, R, C/C++, JavaScript
Data Processing: Big data (Spark), Streaming data (Kafka, Spark Streaming), Orchestration (Airflow)
Certifications: AWS Cloud Practitioner, AWS Solutions Architect Associate
Databases: MySQL, PostgreSQL, Greenplum, Athena, GraphQL
MLOps/DevOps: Docker, Git, Kubernetes, Shell Script
Tools: Machine Learning (PyTorch, TensorFlow, Keras, Scikit-Learn), Pandas, Numpy, PySpark
Data Visualization: Portfolio | **Experienced in**: Customer Segmentation, Churn Prediction, Anomaly Detection
ML Algorithms: Regression (Logistic, Polynomial, Ridge/Lasso), Classification (Logistic, XGBoost, Decision Tree, Random Forest, SVM), Clustering, Bagging, Boosting, Auto-encoders, Deep Learning (CNN, RNN, LSTM)

EXPERIENCE

Data Scientist / Technical Delivery Consultant – Professional Services Team — **Jan 2022 – Sept 2023**
Totogi, Delaware (Remote through Crossover) — Tools: Python, Flask, GraphQL, AWS, Docker, Kubernetes, Linux

- Created a Python ETL tool which accelerated data migration by 30x and migrated over 50 clients
- Demonstrated Totogi capabilities by integrating with Meta's Magma Core on AWS(EKS, S3, EC2) resulting in 3+ new pilots
- Upgraded a legacy software (C/C++), used by 40+ enterprises globally, in 66% less time than expected
- Spearheaded the design and development of web applications and solutions resulting in saving internal team's time by 10 hr/week
- Automated monitoring and testing of Totogi open-source API by creating custom Python tools reducing downtime by 20%
- Delivered 3+ challenging projects out of comfort zone which required learning new technologies (AWS, shell, Flask, Docker)

Data Scientist – Artificial Intelligence (AI) Production — **July 2020 – Jan 2022**
Afiniti, US (Remote in PK Office) — Tools: Python, R, MySQL, Bayesian & Statistical Modeling, Stan

- Increased revenue up to 4% for 5 clients (including Sky BR, Santander MX, ATT MX) by modeling customer-agent behavior using mathematical and machine learning models, ensuring a personalized experience for customers in contact-centers
- Automated fault detection and data integrity using Python and SQL improving downtime by 80% and team's time by 30%
- Aggregated and analyzed 100 GB+ of complex data from 10+ sources to generated over $50,000 additional revenue for clients
- Quantified impact of models with confidence intervals by utilizing statistical analysis (A/B, power, hypothesis tests)
- Designed metrics custom to client's line-of-business to use in revenue optimization and data driven decision making
- Assumed ownership of 5 clients by reviewing and approving data pipelines and models end-to-end prior to deployment
- Collaborated with cross-functional teams from 3 countries to identify business issues and communicated complex analyses
- Supervised 8 data professionals (data engineers, scientist, analysts) and fostered continuous growth and innovation in the team

Research Intern – Processor Architecture Lab (LAP) — **June 2019 – Sept 2019**
EPFL, Switzerland (Prestigious fellowship) — Tools: C++, Verilog, Python, Linux

- Collaborated with 3 EPFL researchers (PhDs and Google Fellow) to benchmark, debug and improve Dynamatic, an open-source dynamically scheduled high-level synthesis tool
- Conceptualized and proposed workarounds after in-depth analysis and investigation of the shortcomings of the tool

PROJECTS

ML Tradestation | Spark, Kafka, Docker, Airflow, MySQL, Python — Github
- A real-time pipeline with Lambda architecture (both Batch and Streaming) to identify trends in stocks/crypto quickly and automatically
- Docker and Docker-Compose to run kafka, airflow, mysql. Airflow for orchestration
- Applied sentiment analysis on news/tweets and trend prediction algorithms on historical data to improve predictions (In progress)

Deep Neural Network on FPGA | C++, Python, PyTorch, Heterogeneous Comp., Linux — Github
- Accelerated Deep Learning inference by 3.36x speedup on FPGA over Intel-i7
- Developed an open source library to accelerate CNNs on FPGA by creating a pipelined architecture of CNNs on FPGA
- Co-optimized hardware and software via restructuring the convolution algorithm and implementing Winograd convolution

Serverless Batch Data Pipeline in Cloud | AWS, Cloudwatch, S3, RDS, AWS Lambda, Python — Diagram
- Automated monitoring and alerts for a production system running on AWS