Abdullah Ashfaq

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

San Diego, CA | (619) 798-8102 | abdullahashfaq11@gmail.com | aashfaq@ucsd.edu

EDUCATION

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

National University of Sciences and Technology (NUST), Islamabad, Pakistan — **Sept 2015 – June 2020**
Bachelor of Electrical Engineering – CGPA: 3.99/4
- Honors: Ranked 2nd out of 180 students. Merit Scholarship for all semesters (Given to top 3 in Class)

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

SKILLS

Certifications: AWS Cloud Practitioner, AWS Solutions Architect Associate | **Databases**: MySQL, Greenplum, Athena, GraphQL
Tools: Airflow, Machine Learning (Pytorch, Tensorflow, Keras, Scikit-Learn), Pandas, Numpy, Heterogeneous Comp., PySpark
ML Algorithms: Regression (Logistic, Polynomial, Ridge/Lasso), Classification (Logistic, XGBoost, Decision Tree, Random Forest, SVM), Clustering, Bagging, Boosting, Auto-encoders, CNN, DNN, RNN, LSTM | **Data Visualization**: Portfolio
Data Science Problems: Customer Segmentation, Churn Prediction, Uplift Modelling (incremental effect)
DevOps: Docker, Git, Kubernetes, Shell Script | **Other.** Communication, Leadership, Teamwork, Critical-thinking, Problem-solving

EXPERIENCE

Software Engineer (Data) – Professional Services Team — **Jan 2022 – Sept 2023**
Totogi, Delaware (Remote) — Used: Python, Flask, GraphQL, AWS, Docker, Kubernetes, Linux

- Accelerated data migration by 30x by designing a Python ETL tool and migrated over 50 clients using it
- Deployed Meta's open-source Magma Core on AWS using Kubernetes, Terraform and Docker and integrated with Totogi OCS
- Upgraded a legacy software (C/C++), used by +40 enterprises globally, in 66% less time than expected (LinkedIn Recommendation)
- Designed, developed, tested tools and services using Python, Flask, AWS to facilitate customers and internal teams
- Automated monitoring and testing of Totogi open-source API by creating custom Python tools
- Delivered challenging projects out of my comfort zone which required learning new technologies (AWS, shell, Flask, Docker)

Data Scientist – Artificial Intelligence (AI) Production — **July 2020 – Jan 2022**
Afiniti, Pakistan (Remote) — Used: Python, R, MySQL, Bayesian & Statistical Modeling

- Increased revenue up to 4% for 5 clients (including Sky BR, Santander MX, ATT MX) through customer retention, segmentation, churn and LTV prediction ensuring a personalized experience for customers in contact-centers
- Ensured data integrity and fault detection via automation decreasing downtime by 80% and saving AI team's time by 30%
- Utilized statistical analysis and testing (A/B, power, hypothesis) for impact quantification with confidence intervals
- Designed metrics custom to client's line-of-business to use in revenue optimization and data driven decision making
- Analyzed Terabytes of complex data to identify optimization opportunities using R and Statistical analysis
- Assumed ownership of clients and projects by reviewing and approving data pipelines and models end-to-end prior to deployment
- Collaborated with cross-functional teams to identify business issues and communicated complex analyses to stakeholders
- 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, Lausanne, Switzerland (Prestigious fellowship) — Used: C++, Verilog, Python, Linux

- Alpha-tester for Dynamatic, an open-source dynamically scheduled high-level synthesis tool
- Investigated the shortcomings of the tool and proposed workarounds after in-depth analyses and experimentation
- Worked with Lana Josipovic (Google Fellow, ETH Zurich) and Andrea Guerrieri on benchmarking and debugging the tool

Research Intern – Machine Learning and AI — **June 2017 - June 2019**
TUKL-NUST R&D Center, Islamabad, PK — Used: Vivado HLS/C++, Python, Pytorch, Heterogeneous Comp., Linux

- Developed open-source library to create custom DL hardware architecture achieving 3.36x speedup on FPGA over Intel-i7
- Restructured algorithms in deep neural networks to achieve Hardware-Software co-optimization
- Implemented the algorithm for binarization using integral image on FPGA

PROJECTS

Anomaly Detection (Github): Used R and Statistics to design dashboard that displays anomalous behavior vs expectation
Serverless Batch ETL Pipeline (Diagram): Used AWS Cloudwatch, S3, Lambda to create serverless pipeline for monitoring
Deep Neural Network on FPGA (Github): Used C++ (HLS) to create flexible library for pipelined dataflow arch. for DNN inference
Self-Balancing Robot (Github): Used Arduino, C and Control Systems theory to create a 2-wheeled self-balancing robot
5 stage Pipelined RISC-V Processor (Github): Used Verilog to write processor and supported S,R and I format instructions