Abdullah Ashfaq

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

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

EDUCATION

University of California San Diego, La Jolla, CA **Expected Graduation: 06-2025**
Master of Science in Data Science

National University of Sciences and Technology (NUST), Islamabad, Pakistan **09-2015 to 06-2020**
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 **01-2018 to 05-2018**
US Dept. of State Global UGRAD Semester Exchange Scholarship

SKILLS

Languages: Python, MySQL, R, C/C++, JavaScript | **Certifications**: AWS Cloud Practitioner, AWS Solutions Architect Associate
Databases: MySQL, Greenplum, Athena, GraphQL | **DevOps**: Docker, Git, Kubernetes, Shell Script
Tools: Airflow, Machine Learning (Pytorch, Tensorflow, Keras, Scikit-Learn), Pandas, Numpy, Heterogeneous Comp., PySpark
Data Visualization: Portfolio | **Experienced in**: Customer Segmentation, Churn Prediction
ML Algorithms: Regression (Logistic, Polynomial, Ridge/Lasso), Classification (Logistic, XGBoost, Decision Tree, Random Forest, SVM), Clustering, Bagging, Boosting, Auto-encoders, CNN, DNN, RNN, LSTM

EXPERIENCE

Data Scientist / Technical Delivery Consultant – Professional Services Team **01-2022 to 09-2023**
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 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 tools 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 **07-2020 to 01-2022**
Afiniti, Pakistan (Remote) 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) **06-2019 to 09-2019**
EPFL, Lausanne, 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

Research Intern – Machine Learning and AI **06-2017 to 06-2019**
TUKL-NUST R&D Center, Islamabad, PK Tools: Vivado HLS/C++, Python, Pytorch, Heterogeneous Comp., Linux

- Accelerated Deep Learning inference by 3.36x speedup on FPGA over Intel-i7 by developing an open-source library to create custom hardware architecture. Achieved Hardware-Software co-optimization via restructuring the convolution algorithm
- Implemented the algorithm for binarization using integral image on FPGA

PROJECTS

Deep Neural Network on FPGA (Github): Developed a flexible library for pipelined dataflow arch. for DNN inference using C++
Anomaly Detection (Github): Designed a dashboard that provided insights resulting in +2% increase in revenue gain
Serverless Batch ETL Pipeline (Diagram): Established a serverless pipeline for monitoring using AWS Cloudwatch, S3, Lambda
Self-Balancing Robot (Github): Built a 2-wheeled self-balancing robot by using Arduino, C and Control Systems theory
5 stage Pipelined RISC-V Processor (Github): Created a processor which supported S,R and I format instructions using Verilog