Nishat Khan

nishatconnect@gmail.com | 628.502.3712
linkedin.com/in/nishat-khan1 | nishat-khan.github.io | github.com/nishat-khan

Professional Experience

LeanTaaS   **Jul 2021-present**

Senior Data Engineer
- Designed and implemented the data pipeline for a real-time hospital bed capacity management system.
- Led the data pipeline team for 8 months following the departure of 2 senior software engineers and a system architect and streamlined the onboarding process for 3 new engineers.
- Achieved USD 1.5 million in annual recurring revenue by expanding to 2 new hospital networks.
- Mentored a team of 5 engineers and provided code review, approval, and production troubleshooting support
- Experience with Python, AWS services, Airflow, GraphQL, Jenkins.

Pure Storage   **Oct 2020-Jul 2021**

Data Analyst, Bay Area
- Automated rightsizing recommendations of EC2, RDS, EBS instances on AWS by analyzing key metrics and usage data, saved half a million USD annually.

Checkpoint Software Technologies   **Jan 2020-Jun 2020**

Data Science Intern, Bay Area
- Utilized machine learning techniques to improve the cybersecurity of cloud serverless entities.
- Improved unsupervised anomaly detection models by clustering temporal data using an ensemble of k-means and Isolation Forests, resulting in a significant reduction in false positives (from 25% to 0.07%).
- Synthesized weakly labeled data using SMOTE to validate model performance and achieved an F1 score of 80%.

Honeywell Technology Solutions

Tech Lead, Bangalore   **Jul 2016-Jun 2018**
- Reduced the purchase cost of licenses by comparing ARIMA time series and regression models in Python to predict license-based software usage for all engineering operations.
- Designed a chatbot using Long Short Term Memory Model (LSTM) for proprietary software installation and bug reporting.

Senior Engineer, Hyderabad   **Jun 2018-Jun 2019**
- Created a Django and SQL application to lower the cost of Satcom1's legacy billing system (20M USD revenue per year) and integrated it with the billing portal, resulting in cost savings for the startup (which was later acquired by Honeywell).
- Improved the flight crew alerting system by replacing a heuristic-based approach with a Random Forest algorithm, resulting in a 70% reduction in CPU usage and increased code maintainability.
- Optimized cockpit information display by redesigning the user interface of 4 monitors in C++ and consolidating it onto 3 monitors.

Teaching Assistant for Machine and Deep Learning, University of San Francisco
- Computation for Analytics in Python, Advanced Machine Learning in PyTorch   **Aug 2020-May 2021**

Education

University of San Francisco   **Jul 2019-Jun 2020**

Master of Science in Data Science
Courses: Machine Learning, Deep Learning, Time Series Analysis, Distributed Computing, A/B Testing.

IIT Hyderabad   **Aug 2012-May 2016**
B.Tech. in Electrical Engineering with Honors in Image Processing
Courses: Linear Algebra, Probability, Machine Learning and Data Structures and Algorithms

Research Paper

Automated eye diagnosis of cataract patients (~5000 per day) with an image segmentation software using patch-based Particle Swarm Optimization in MATLAB. Paper published in *European Journal of Ophthalmology*.

Technical Skills

Languages: Python, SQL, Javascript
Python Libraries: Scikit-Learn, NumPy, Pandas, PyTorch, PySpark, Plotly, Seaborn
Technologies: AWS, Airflow, Git, Jenkins, GraphQL