

Fahim Rustamy

Route 212, Richlandtown, PA, 18955 | (610) 653 3925 | faheemrustamy@gmail.com | <https://rustamyf.github.io>

SUMMARY

Seeking a full-time position in data science and machine learning.

Engineering Ph.D. with a focus on probabilistic modeling and machine learning with 5+ years of professional and academic experience. Big data and cloud computing is my passion. 4X AWS Certified

EXPERIENCE

ATLSS Center, Lehigh University

August 2017 - Present

Data Scientist/Probabilistic Modeler/Research Assistant, Bethlehem, PA

- Developed synthetic seismic ground motion records using stochastic processes for the Eastern US. Processed, cleaned and organized over 52 GB of ground motion data with Matlab and Python.
- Proposed a new method for evaluating the fragility of engineering structures using Artificial Neural Networks that reduces the computational time by 70% from the traditional method.
- Developed system reliability of power systems using Monte Carlo Simulations and Logistic Regression Methods.
- Proposed a simulation method using damage data and multivariate random samples for system resilience.

State Corps

October 2012 - May 2015

Project Manager, Kabul, Afghanistan

- Managed \$30 million worth of infrastructure projects and supervised over 500 employees.
- Completed the projects on time and was awarded several completion awards from USACE & the government.

PROJECTS

ML and Cloud:

- Fine-tuned a ResNet model using Tensorflow for object recognition, and an NLP model using Spacy to extract information from text. Deployed them to AWS lambda and created an API for prediction.
- Trained the Boston Housing data using logistic regression, support vector machine, and XGBoost. Created a rest API and a web app using Flask. Packed the application with Docker and deployed it with AWS ECS.
- Created a Q&A NLP model using BERT and HuggingFace transformer. Deployed a rest API using flask.

Big Data and Cloud:

- Prepared, cleaned, and analyzed 135 million Amazon reviews (35 GB) using Athena/Presto and Spark. Trained a classifier using Tensorflow framework and HuggingFace transformer with 93 % accuracy that takes the text review as input and classifies its rating from 1 to 5. Deployed the model using AWS Sagemaker.
- Performed exploratory analysis of NYC 50 million taxi trips with ~200 GP of data using PySpark, Spark SQL and AWS MapReduce.

EDUCATION

Lehigh University, Bethlehem, PA

August 2015 - May 2021

Ph.D. & M.Sc. in Engineering/Probabilistic Modeling of Engineering Structures

- Relevant Coursework: Deep-learning, Machine Learning, Random Vibration, Engineering Probability, Data Mining, Engineering Risk, Optimization Theory and Time Series Analysis

Kabul University, Kabul, Afghanistan

March 2009 - January 2013

B.Sc. in Civil Engineering

SKILLS

Languages	SQL, Python (Pandas, Numpy, Matplotlib, scikit-learn, Flask and Tensorow), and Matlab.
Cloud	Sagemaker, Athena, Redshift, S3, EMR, Hadoop, AWS Lambda, Kubernetes and other AWS tools.
ML Models	Linear and Logistic Regression, XGBoost, Neural Network, RNN, CNN, and KNN
Others	LaTex, Docker, GIT and Linux

CERTIFICATES

- AWS Developer Associate
- AWS Machine Learning Specialty
- AWS Cloud Practitioner
- AWS Data Analytics Specialty
- Tensorflow Developer Certificate