Muhammad Maaz Kamal

+1 929-409-9885
mukamal@syr.edu
New York

linkedin.com/in/maazkamal/
github.com/mukamal/
mukamal.github.io/

Education

M.S. Applied Data Science - Syracuse University, School of Information Studies May 2021
Coursework: Big Data Analytics, Natural Language Processing, Text Mining, SQL Data Management, Linear Statistical Models, Financial Analytics, Business Analytics

B.S. Electrical Engineering - Lahore University of Management Sciences May 2016
Coursework: Computer Networks, Digital Communication Principles, Data Structures, and Principles of Economics

Skills

Programming Languages:	Python, Java, R, C++, MATLAB, Assembly Language
Machine Learning:	Classification, Regression, Association, Clustering, Time Series
Machine Learning Algorithms:	Decision Tree, Random Forest, Gradient Boosting Machine, Neural Network, Linear and Logistic Regression, Naïve Bayes, K- Nearest Neighbors, Association Rule Mining, Support Vector Machines
Database Management:	SQL, MySQL, Microsoft Access
Data Analysis:	Hypothesis Testing, TensorFlow, PySpark, Pandas, Scikit, dplyr, Google Analytics, Microsoft Excel
Data Visualization:	Tableau, Power BI, Amazon QuickSight, Matplotlib, ggplot, Microsoft PowerPoint

Projects

Risk Assessment of Loan Investment Dec 2020
- Analyzed large dataset (150+ features and 2.5 million+ records) using distributed computing to evaluate applicants' financial health by building Random Forest and Gradient Boosted Tree models.
- Utilizing PySpark and SciPy, built a model that could predict whether an individual defaults on loan, with an accuracy of 77%- which was an improvement of 33% compared to majority class baseline.

Text Summarization Dec 2020
- Utilizing TensorFlow and NLTK, performed Seq2Seq modelling on a dataset of 100,00 news articles to build a text-summarization model that could summarize any given (English) article.
- The model summarizations were evaluated using Bilingual Evaluation Understudy (BLUE) score- the resulting cumulative score turned out to be 0.6 (where a score of 1 means a perfect match and a score of 0 means no match).

Recommender System Oct 2020
- Analyzed dataset containing all Shakespeare quotes of 966 characters using Principal Component Analysis (PCA) to build a Recommender System that makes recommendations on new characters to study based on a specific character that a Shakespeare fan likes.
- The computational cost of training machine learning model was significantly reduced by choosing only 150 principal components (which still captured 90% of the variance in the data).

Work Experience

Data Analyst- Intern, Guest Worldwide July 2020- Aug 2020
- Improved critical insights and access of senior decision makers by automating the analytics pipeline and creating interactive dashboards and reports, in Tableau, for all 14 subsidiaries of Guest Worldwide.
- Preprocessed and prepared data acquired from disparate sources and enhanced the management and maintenance of information system using Excel VBA.

Technical Project Manager, Khud Initiative Apr 2017 – Jul 2019
- Supervised 3 tech startups by keeping track of KPIs and providing marketing recommendations by analyzing data from Google Analytics, A/B testing and digital campaigns.
- Led team of 10+ and communicated project plans and progress to key stakeholders, including business and technical resources.

CEO/ Project Manager, Interacta Jul 2014 – Apr 2017
- Founded Interacta, built MVP and secured investment of $ 220,000 (USD) based on the MVP.
- Led team of 4 to develop core algorithm and optimization of audio recognition on the backend using Python Django.
- Led team of 2 to design SQL database to support the algorithm, keeping optimization in view, and scripted SQL queries.