

Vishwesh Srinivasan

vishweshsrini@gmail.com | 781-921-4385 | [LinkedIn](#) | [GitHub](#) | [Tableau](#)

EDUCATION

Tufts University

Master of Science in Data Analytics | GPA: 3.96

Coursework: Machine Learning, Big Data, Data Science for Urban Sustainability, Database Design & SQL, Adv. Statistics I & II

Teaching Assistant: Foundation of Data Analytics (DATA200) & Introduction to Data Analytics (DATA100)

Medford, MA, USA

09/2022 – 05/2024

National Institute of Technology Warangal

Bachelor of Technology in Mechanical Engineering

Warangal, India

08/2016 – 08/2020

EXPERIENCE

Paragon Corporation, MA, USA

Data Analyst

Remote, USA

01/2025 – Present

- Improved the efficiency of the ETL process for the payroll data of 25+ clients by streamlining the data extraction from Excel/PDF files, standardizing formats, and loading data into a PostgreSQL database using Python and SQL scripts.

Success Academy Charter Schools

Data Analyst

New York, NY, USA

07/2024 – 01/2025

- Reduced feedback turnaround time by 85%, enabling faster teaching adjustments and improving student outcomes by automating the generation of academic reports in Google Sheets, delivered to the school leadership and teachers.

Data Intensive Studies Center, Tufts University

Research Intern ([Poster](#))

Medford, MA, USA

01/2024 – 05/2024

- Developed a novel approach to detect rare tumor cell clusters in blood by training a random forest classifier using a correlation-based feature-engineered light scattering data (Accuracy, Precision, TPR, TNR: ~90%).

JPMorgan Chase & Co.

Data Science Intern, Credit Risk Modeling

Wilmington, DE, USA

06/2023 – 09/2023

- Developed two XGBoost classifiers (with and without additional guarantors' indicators) using a dataset of 1M+ records for small business credit decision-making and analyzed the models' performance across different data segments.
- Leveraged the analysis to propose a strategy reducing the documentation and approval time by 30% (est.) for applications with less than \$250K exposure by eliminating the need for additional guarantors' information.
- Facilitated the migration of PySpark-based model development from local infrastructure to the AWS ecosystem to reduce costs by collaborating with the infrastructure team and establishing the best practices.

Citigroup

Software Developer, Personal Banking & Wealth Management Technology

Chennai, India

08/2020 – 09/2022

- Simplified the personal loan processing workflows by re-engineering the end-to-end statement generation and payment hierarchy application logic, working within a cross-functional team.
- Automated report processing (saving 10+ hrs/week of manual effort) by developing a VBA system that converts text reports into dashboards, delivering insights on product metrics like new account activity, balances, and transaction volumes.

Language Technologies Research Center, IIIT Hyderabad

Research Intern ([Repository](#))

Hyderabad, India

05/2019 – 07/2019

- Reduced the perplexity by 16% compared to the SOTA sentence simplification model by training a Seq2Seq model using the WIKISPLIT dataset with reward (BLEU score) augmented maximum likelihood objective function.

Reliance Jio Infocomm Limited

Data Science Intern

Mumbai, India

05/2018 – 07/2018

- Boosted the search efficiency of an internal platform used for resolving network coverage issues by building related search (NCD algorithm), auto-complete (N-Grams), and recommendation (TF-IDF & cosine similarity) features.

SPI Cinemas Private Limited

Data Science Intern

Chennai, India

11/2017 – 12/2017

- Decreased the turnover rate by 30% (est.) by pioneering a data-driven hiring strategy and developing a predictive model (87% accuracy) to predict the likelihood of frontline staff leaving within the first three months.

TECHNICAL SKILLS

Programming Languages: Python, R, SQL, MATLAB, C++

Competencies: Linear & Logistic Regression, Tree-Based Methods, Ensemble Learning, Neural Networks, CNNs, Text & XML Processing, Word Embeddings, LSTM, Hypothesis Testing, SHAP, Partial Dependence Plots

Tools/Libraries: GCP (BigQuery), AWS (S3, SageMaker, EMR), Tableau, Power BI, R Shiny, NumPy, Pandas, Matplotlib, Scikit-Learn, PySpark, GeoPandas, PySAL, PyTorch, TensorFlow, ggplot2, Linux, Git, Jira