

# Vishwesh Srinivasan

[vishweshsrini@gmail.com](mailto:vishweshsrini@gmail.com) | +1 (781) 921-4385 | [LinkedIn](#) | [GitHub](#) | [Tableau](#)

## EDUCATION

**Tufts University, MA, USA** | Master of Science in Data Analytics | GPA: 3.96/4.00 Sep 2022 – May 2024  
**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)  
**National Institute of Technology Warangal, India** | B.Tech in Mechanical Engineering Aug 2016 – Aug 2020

## EXPERIENCE

**Paragon Corporation, MA, USA** | Database Programmer Jan 2024 – May 2024; Jan 2025 - Present

- Increasing the efficiency of clients' payroll systems by streamlining the maintenance of their PostgreSQL databases.

**Success Academy Charter Schools, NY, USA** | Data Associate July 2024 – Jan 2025

- Reduced feedback turnaround time by 85%, enabling quicker teaching adjustments and enhancing student outcomes by streamlining academic data management and delivering daily assessment reports in Google Sheets to leadership.

**Data Intensive Studies Center, Tufts University, MA, USA** | Research Intern ([Poster](#)) Jan 2024 – May 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., DE, USA** | Applied AI ML Sr. Associate, Credit Risk Modeling Jun 2023 – Sep 2023

- Reduced the documentation and approval time for small business customers (below \$250k exposure) by 30% (est.) by reducing required data points to only the primary guarantor's indicators, thereby improving customer experience.
- Formulated the above strategy by building and analyzing the performance of two XGBoost classifiers on various data segments using PySpark on AWS, with and without additional guarantor indicators.

**Data Analytics Department, Tufts University, MA, USA** | Graduate Research Assistant ([Website](#)) Jan 2023 – Apr 2023

- Automated 80% of the name entity tagging process of D'Arcy Thompson's Glossary of Greek Birds book (second edition) to make it accessible to the general audience using XML text processing and NLP techniques.

**Citigroup, India** | Software Developer, Personal Banking & Wealth Management Technology Aug 2020 – Jul 2022

- Increased the efficiency of personal loan processing by simplifying the workflows in the core banking systems.
- Eliminated the manual processing of reports with an automated system to convert them to dashboards using VBA.

**Language Technologies Research Center, IIIT Hyderabad, India** | Research Intern ([Repository](#)) May 2019 – Jul 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, India** | Machine Learning Intern May 2018 – Jul 2018

- Improved 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, India** | Data Science Intern Nov 2017 – Dec 2017

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

## PROJECTS

**Relationship between air pollution and walkability index in Greater Boston Region** ([Repository](#))

- Found a positive correlation between exposure to air pollution and the walkability index by analyzing the visualizations (using GeoPandas) and the results of spatial regression models implemented using PySAL.

**Gentrification study of New York and Los Angeles metropolitan areas** ([Repository](#))

- Implemented classification algorithms with SMOTE techniques (to handle imbalanced data) to predict the likelihood of a census tract getting gentrified between 2000 and 2010 using the Neighborhood Change Database.

## TECHNICAL SKILLS

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

**Programming Languages:** Python, R, SQL, MATLAB, Visual Basic for Applications (VBA), C++

**Tools:** GCP, AWS, Tableau, NumPy, Pandas, Matplotlib, Sklearn, PySpark, PyTorch, TensorFlow, ggplot2, R Shiny, Linux, Git