

Vishwesh Srinivasan (Open to work anywhere in the U.S.)

Vishwesh.Srinivasan@tufts.edu | +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: Database Design & SQL, Intro to Machine Learning, Big Data, Data Science for Urban Sustainability, Adv. Statistics I & II
Experience: Graduate Teaching Assistant for 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*

TECHNICAL SKILLS

Programming Languages: Python, R, SQL, MATLAB, Visual Basic for Applications (VBA), C++
Tools: GCP, AWS SageMaker & EMR, Tableau, R Shiny, Elasticsearch, Kibana, Kepler.gl, Linux, Git, Jira, Excel, PowerPoint
Libraries: NumPy, Pandas, Matplotlib, Scikit-Learn, PySpark, GeoPandas, PySAL, PyTorch, TensorFlow, ggplot

EXPERIENCE

- Data Intensive Studies Center, Tufts University, MA, USA** | Research Intern *Jan 2024 – May 2024*
- Enabled label-free flow cytometry of rare circulating tumor cell clusters in blood with correlation-based feature engineering of light scattering data and a random forest classifier (Accuracy, Precision, TPR, TNR: ~90%). ([Poster](#))
- Paragon Corporation, MA, USA** | Database Programmer *Jan 2024 – May 2024*
- Increased the efficiency of clients' payroll operations by streamlining the development and maintenance of their PostgreSQL databases.
- JPMorgan Chase & Co., DE, USA** | AI & Data Science Summer Associate, CCB Risk Modeling *Jun 2023 – Sep 2023*
- Enabled faster credit approvals and improved customer experience for small businesses by limiting the data points required from customers with less than \$250k exposure to only the primary guarantor's indicators.
 - Formulated the above strategy by building and analyzing the performance of two XGBoost classifiers using PySpark on the AWS cloud, with and without the additional guarantors' indicators.
- Data Analytics Department, Tufts University, MA, USA** | Graduate Research Assistant *Jan 2023 – Apr 2023*
- Contributed to making D'Arcy Thompson's Glossary of Greek Birds accessible to the general audience using automated tagging and natural language processing techniques.
- Citigroup, India** | Tech Program Application Developer – 1, PBWM Technology *Aug 2020 – Jul 2022*
- Enabled the efficient processing of loans by streamlining and simplifying the processes 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 *May 2019 – Jul 2019*
- Reduced the model's perplexity used for sentence simplification by 16% compared to the state-of-the-art model by implementing a Seq2Seq model with a reward function in Python.
- Reliance Jio Infocomm Limited, India** | Machine Learning Intern, Jio Coverage Platform *May 2018 – Jul 2018*
- Enhanced the search experience of an internal platform, used to monitor and fix network coverage issues by building auto-complete, related search features, and a recommendation system using NLP algorithms in R.
- SPI Cinemas Private Limited, India** | Data Science Intern, Human Resources *Nov 2017 – Dec 2017*
- Pioneered a data-driven hiring culture by developing a pipeline to visualize data and a classification model with 87% accuracy in Python to predict the likelihood of a frontline employee leaving within the first three months of joining.

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 in the Greater Boston Region by analyzing the visualizations and the results of spatial regression models.
- Database system to manage the payroll system at Tufts Dining** ([Repository](#))
- Designed a database system and developed a user guide to manage the database. The user guide contains queries to add data and the most frequent scenarios for updating, deleting, and viewing the data in different forms.
- 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.