VISHESH GUPTA

📞 857-313-0666 | Boston, MA, 02115 | ✉ gupta.vishe@northeastern.edu | 💼 linkedin | ⬛ Github | 🌐 Website

EDUCATION

Northeastern University (Boston, USA)   Expected May'24
Candidate for Master of Science in Data Analytics Engineering, **GPA: 4.0/4.0**
Relevant Coursework: Foundations of Data Analytics, Data Management for Analytics

SRM University (Chennai, India)   Jun'17 – May'21
Bachelor of Technology in Information Technology, **GPA: 3.54/4.0**
Relevant Coursework: Data Science and Big Data Analytics, Visualization Techniques, Statistics, Probability & Queuing Theory, Discrete Mathematics, Database Management Systems

TECHNICAL SKILLS

Languages: SQL, Python (NumPy, Pandas, Seaborn, Matplotlib, Scikit-learn), R, Flask
Statistical & Analytics: Data Analysis, Data Visualization, Data Mining, Classification, Clustering, Regression, NLP
Framework and Databases: Tableau, Microsoft Power BI, RStudio, Putty, MySQL, PostgreSQL, PyCharm, Git, Excel

EXPERIENCE

Tata Consultancy Services, New Delhi, India: System Engineer   May'21 – Jun'22
- Performed data exploration and manipulation on Grafana model using correlation to create dashboards and actionable insights using Power BI, resulting in a **17% increase in productivity**
- Extracted, cleansed, and loaded datasets from copies of production data, data warehouses, and other sources and incorporated descriptive data analysis and regression analysis to propose business strategies
- Conducted internal/external validation methods on product survey data to determine prediction model's accuracy
- Performed data manipulation on a workflow for periodic migration of data from DynamoDB to PostgreSQL
- Utilized tools such as, Bitbucket, Jenkins, Redis Cache, Grafana to engage with complex data sets for predictive modeling and building dashboards to show KPI metrics which decreased the **client server cancellations by 5%**

DLK Technologies, Chennai, India: IT-Analyst (Intern)   Dec'19 – Jan'20
- Created and produced visualizations of multivariate categorical data in R using ggplot2 to depict deployment pattern
- Optimized and enhanced existing SAP Business Object dashboards to perform year over year, quarter over quarter, YTD, QTD, and MTD types of analysis to gain a better understanding of complex data root failure parameters

Vocable Private Limited, Gurugram, India: IT-Operations (Intern)   May'19 – Jun'19
- Developed a VBA tool to automate the daily task of sending reports to clients thereby reducing the time taken from **4 hours daily to 10 minutes**, achieving full accuracy while saving **650 hours/year**
- Analyzed marketing campaigns using Test Control methodology and set up multiple databases on MS SQL Server; generated key insights on Google Analytics assisting Vocable clients in capturing 10% more resource allocation

PROJECTS

Truth Detection on Twitter Dallas Dataset, Northeastern University   Sep'22 – Jan'23
- Created a new Scalable Robust Truth Discovery (SRTD) model that takes into account multiple source attitudes, claim quality analysis in a systematic truth discovery approach in tweets
- Explicitly calculated the source reliability, report credibility, and a source's historical behaviors to effectively address the misinformation spread and data sparsity challenges, resulting in the **truthfulness scores of 80% of tweets**
- Designed and implemented a distributed framework using Work-Queue and the HTCondor system to handle the scalability challenge of the problem, providing 19% of the tweets claims in the 2016 Dallas shooting data trace

Analysis on Forecasting System for Coronary Heart Disorder, Northeastern University   Oct'22 – Dec'22
- Combined 7 datasets with **20+ million rows and 200+ columns**, performed exploratory data analysis to prepare data for modeling, and proposed a Web-Based Portal to predict individual heart disorders
- Implemented full Data Science pipeline to build a feature selection and classification model to predict the least amount attributes with maximum accuracy using ensemble learning methods, resulting in best prediction accuracy of 93.7%

RESEARCH PUBLICATIONS

Springer Book Series "Algorithms for Intelligent Systems"   Jan'22
- Published Paper - Truth Detection Algorithm in Social Media tweets using Similarity measures

International Journal of Recent Technology and Engineering   Mar'20
- Published Paper - The Significance of Predictive Analysis in Predicting Organizational Members' Performance