Tushar Srivastava
Boston, MA | (781) 824-6789 | tushar98@bu.edu | linkedIn.com/in/tusharsri1998

EDUCATION

Master of Science, Computer Information Systems (Data Analytics Concentration) Expected May 2023
Boston University, Massachusetts, USA
Relevant Courses - Data Science with Python, Advanced Database Management, Foundation of Analytics with R, Data Visualization with R, Web Analytics Mining, Advance Machine Learning.

Bachelor of Technology, Computer Science August 2016 - May 2020 IIITDM, Jabalpur, India

SKILLS/TOOLS

Programming Languages: Python (NumPy, Pandas, Seaborn, Scikit-Learn, Keras), R (dplyr, tidyverse), SQL, JavaScript
Databases: PostgreSQL, MongoDB, SQL Server, MySQL
Data Visualization: Google Data Studio, Tableau, Excel, PowerBI
ML Algorithms: KNN, Logistic Regression, Random Forest, Naïve Bayes, K-means, Ensemble, Support Vector Machine,
Tools/Frameworks: Visual Studio, Anaconda Navigator, R Studio, MongoDB Compass, Django, Flask, MEAN Stack, AWS

WORK EXPERIENCE

Data Science Research Assistant - Boston University, Boston, MA August 2021 - Present
- Collected data from SEERSTATS database of breast cancer patients, performed data cleaning, processing, and analysis of MassGeneral hospital patients with survival record of 5 years and 10 years based on various treatment procedures.
- Currently working under guidance of Dr. Saveli Goldberg and Prof. Eugene Pinsky to find the best classifier (Logistic 81% accuracy, Random Forest 84%, Neural Network 87% etc.) hyperparameter tuning to classify survival status.
- The end goal is to provide an optimal treatment procedure for survival based on patient's medical history.

Teaching Assistant - Boston University, Boston, MA August 2021 - December 2021
- Worked as teaching assistant for both Information structures with python and foundation of analytics with R courses.
- Responsibilities included taking informative sessions for a batch of 60 students and helping professor with grading.

Data Analyst Intern, Quantiphi Inc, Mumbai, India August 2020 - January 2021
- Acquired hands-on experience in various tools and technologies namely Anaconda, Angular 8, Node, Python, Django, AWS Cloud services, MongoDB.
- Worked with AFLAC Clients (US) in developing a cloud-based (AWS) solution to fetch data from On-premises database, clean it and load it to Amazon RDS and send to datarobot for data analysis.

Software Development Intern, Smartivity Labs, Delhi, India May 2019 - November 2019
- Led a team to design a Django application for retaining track of Employee's daily work and generate a report showing funds spent on each project. Deployed it on IIS Server in AWS EC2.
- Developed a Process Flow Automation Software to keep track of progress of all product development phases utilizing MEAN Stack and devised an efficient algorithm to provide optimal product packaging combination.

ACADEMIC PROJECTS

Data Warehouse using ETL and Web Scraping September 2021 - December 2021
- Analyzed tennis game statistics to study various player performances based on parameters like court type, player stats, tournaments. Followed ETL for data wrangling.
- Scraped and preprocessed data from multiple sources using python scripts, Transformed and Loaded data in PostgreSQL.
- Designed a dashboard in Google Data Studio to visualize player stats trend and optimized query runtime by 12%.

GIF generation using Posenet Pose detection data and Clustering September 2021 - December 2021
- Used Posenet output tensor data for 130 different physical exercises, decomposed the tensor using explode method and generated GIF.
- Utilized methods like PCA, t-SNE and u-map to reduce dimensions to 2, obtained 4 optimal number of clusters utilizing elbow method, implemented clustering methods like K-Means, DBScan, Birch and visualized 4 clusters in 2D and 3D.

Data Analysis of Tennis Games and Winner Prediction May 2021 - August 2021
- Collected data of all tennis matches from year 2015 to 2019, cleaned, processed, and analyzed data based on various features to create a functional dataset using NumPy and Pandas.
- Used Various Machine Learning Classifiers to find best model, achieved accuracy of 92% using Logistic Regression for winner prediction and plot trends using Matplotlib.
- Developed a user interface using Flask to predict the winner and display player stats.