RUBAL SHARMA

Phone: 315-278-2862 | Email: rubal1999@gmail.com | www.linkedin.com/in/rubal-sharma99/

EDUCATION

Syracuse University- Syracuse, NY (GPA: 3.8) **May 2023**
Master of Science in Computer Science

S.R.M. Institute of Science and Technology- Chennai, India (GPA: 8.8)
Bachelor of Technology in Computer Science and Engineering **May 2021**

TECHNICAL SKILLS

- **Languages**: C, C++, Java, R, Python, HTML, CSS, PHP, SQL
- **Tools**: Jupyter NB, Excel, Google Collab, R Studio, MS SQL server, OracleDB, MongoDB, Azure, GCP, AWS, PowerBI, Tableau
- **Niche skills**: Database Management- SQL & XML databases; Machine Learning- data modelling, neural networks & NLP; Data Science- extracting & pre-processing data, A/B testing, exploratory data analysis, data visualization, statistical analysis.

EXPERIENCE

Data Science Intern, Syracuse University, Syracuse (NY) **July 2023- Present**

- Conducted in-depth research on stock market prediction, leveraging historical data, financial indicators & ML models.
- Developed and fine-tuned an ensemble predictive model, combining Support Vector Regression (SVR) and Gradient Boosting Machines (GBM) to achieve a high level of accuracy in forecasting stock prices. Utilized Git for code tracking & collaboration.
- Conducted comprehensive sensitivity analysis to check the model's robustness across diverse market scenarios & economic conditions. Collaborated with domain experts to validate model predictions and gather critical insights into market dynamics.

Executive, The Climber, Chennai **February 2019- December 2019**

- Organized 2 events Kickstart & Youth Conclave with more than 500 participants. Developed communication & social skills.
- Designed an Excel-based financial analysis and budgeting tool, automating data integration, complex calculations, and scenario analysis, resulting in a 50% reduction in manual processing time.

Software Engineering Intern, Uniq Technologies, Chennai **May 2019- August 2019**

- Built recommendation engines (python) with development team. Worked on websites on Flask, HTML & CSS for 5 clients.
- Aided in migration of data from SQL Server to Cloud Database through Azure Synapse and SQL DB. Analyzed & visualized demographics of user data using PowerBI. Leveraged PowerBI tools, specifically maps, line & scatter plots for weekly reports.

RESEARCH PUBLICATIONS

A comprehensive survey on Free Parking Space, Road Signs and Lane Detection. **December 2021**

- Implemented a new interface to combine detection of Free Parking Space, Road Signs and Lane detection using a Convolutional Neural Network on Google Collaboratory and Spark. Used and compared VGG, OpenCV, TensorFlow, Keras to classify the images.

Maize Leaf Diseases Classification Using Densely Connected Convolutional Network. **November 2021**

- Achieved 98.45% of accuracy on detecting the plant diseases on Maize leaf images and optimized model on noisy (blurry) images to get 91% accuracy which proved the robustness. Formulated neural network using DenseNet121 on Maize leaf images.

Deep Model for Robust Tomato Disease Detection on Low-Resolution Leaf Images. **October 2021**

- Detected diseases in tomato leaves with 99.34% accuracy. Pre-processed & analyzed plant village dataset of 50,000 images. Applied transfer learning on ResNet101 and classified 10 classes including healthy class and 9 disease classes of tomato plants.

PROJECTS

Predicting Employee Retention for Sailfort Motors **July 2023-August 2023**

- Predicted retention by 96.7% accuracy. Evaluated logistic regression & random forest using AUC, F-1 score, precision & recall.
- Used NumPy, Pandas for ETL process. Employed matplotlib & seaborn for visualization to gain insights into the data.

Recruitment Management Database **December 2022- January 2023**

- Designed a database of more than 20 tables using Vertabelo. Created keys and relationships between tables. Used Hadoop to store.
- Deployed tabular design on MySQL Server. Wrote scripts in SQL and inserted data. Utilized Visual Studio for reports.

Prediction of Credit Card Spending **July 2022- August 2022**

- Predicted credit card spending of a user with more than 90% accuracy on R studio. After pre-processing data, implemented and compared random forest, Naïve Bayes & Linear Regression for training and testing using Scikit-learn.
- Envisioned data of various users on PowerBI to understand driving factors for credit card spending.

Dashboard for Cryptocurrency **June 2022- July 2022**

- Designed the website by HTML & CSS to display a dashboard leveraging PowerBI; showing value, statistics and fluctuations over days, weeks and months of 20 Cryptocurrencies through scatter plots, histograms and waterfall charts.
- Built a database Using MyPHP Admin and used AJAX to collect crypto data daily and store it in database.

Fan Favorite prediction and Demographic analysis of Formula 1 using Tweets **September 2021- December 2021**

- Collected tweets using twitter search API, tokenized data, removed stop words using NLTK & stored in MongoDB. Deployed VADER to analyze sentiments, classified tweets as positive, negative & neutral and predicted fan favorite driver.
- Utilized Tableau to visualize demographics of drivers & corresponding tweets from 20 different countries through plots & charts.

CERTIFICATIONS

- Google Advanced Data Analytics Professional Certificate | Coursera
- Machine Learning, Stanford University | Coursera
- Data Science using Python | NPTEL