SIMHAVISHNU RAM PRASAD

simhavishnurp@gmail.com| (716)-650-8642 | https://www.linkedin.com/in/simhavishnuramprasad/ | https://github.com/Simhavrp

EDUCATION

University at Buffalo, Buffalo – NY, USA *Aug 2022– May 2024*
Masters in Engineering Science: Data Science **GPA:3.23**
Relevant Courses: Database Modelling and Query Languages, Statistical Learning, Probability, Numerical Mathematics for Data Scientists, Machine Learning, Introduction to Time Series, Predictive Analytics

National Institute of Technology, Durgapur – West Bengal, India *Aug 2018– May2022*
Bachelors in Computer Science and Engineering **GPA:7.52/10**
Relevant Courses: Design and Analysis of Algorithms, Data Structures and Algorithms, Machine Learning, Object Oriented Programming, Artificial Intelligence

SKILLS

- Technical Languages: Python, C/C++, R, SQL, Java, Dart, MATLAB, Octave, Julia
- Cloud & Hosting Platforms: AWS, Azure, Heroku, Docker, CI/CD Pipelines
- Frameworks: PyTorch, TensorFlow, HuggingFace, SciPy, NLTK, NumPy, Matplotlib Pandas, OpenCV, Angular, Express, Flask
- Tools: Firebase, RESTAPI, JSON, Tableau, PowerBI, Git/GitHub, Anaconda, R Studio, Jupyter Notebook, Docker, Latex, Linux, Postgres
- Industry Knowledge: Generative AI, Language Models, Deep Learning, Computer Vision, Natural Language Processing, Android

EXPERIENCE

Data Science Research Intern, *Council for Scientific and Industrial Research* *Feb 2021 – May 2021*

- Developed a novel weighted voting-based ensemble algorithm integrating visual (CNN) and sensor-based (Random Forest) classifiers, for enhanced fire detection; achieved 97.58% accuracy.
- Noticed a 60% reduction in false alarm rates while maintaining swift and precise fire detection.
- Lead-authored a paper titled " Optimized Weighted Ensemble Voting Based Machine Learning Framework for Improved Fire Detection Using Longitudinal Multi-Sensor Data" under review at the IETE Journal of Research. (Link)

Data Analyst Intern, *Tata Consultancy Services* *Nov 2020 – Jan 2021*

- Analyzed net gross margin, monthly sales growth, and website traffic using R, identifying key trends that contributed to a 12% increase in quarterly sales.
- Created interactive dashboards and visualizations in QlikSense to track key performance indicators, such as sales revenue and customer retention rates, reducing report generation time by 20%.
- Used SQL queries, joins, and aggregations to manipulate and organize databases logically, ensuring data integrity and accessibility for users.
- Analyzed customer behavior through cohort analysis and segmentation, identifying retention trends that led to a 5% increase in repeat purchase rates over two months.

PROJECTS

- **New York Farmer's Market Analysis** (Github Link): Developed a data-driven web application analyzing the New York Farmers' Market Dataset, leveraging Spring Boot, Spring Data JPA for backend, Streamlit for UI, integrated SQL database management, and deployment.
- **Classification of EphA4 Receptor Antagonist Inhibitors** (Github Link): Developed a machine learning model using XGBoost to classify EphA4 receptor antagonist inhibitors from the AID 689 dataset, facilitating therapeutic research for spinal cord injuries, thrombosis, and cancer treatments; achieved test accuracy of 93%.
- **UFC Data Analysis with AWS Redshift**: Leveraged AWS Redshift to build and maintain a UFC SQL database, analyzing correlations between fight outcomes and fighter performance. Designed an ETL pipeline to process over 30,000 data points, ensuring data quality through normalization, deduplication, and cleanup techniques.
- **Study of High-Energy Particle Discrimination Using Classification Algorithms** (Github Link): Did a comparative study of classification models for discriminating between primary gamma signals and hadronic showers initiated by cosmic rays using various classification algorithms like Decision Trees, AdaBoost, K-NN, Logistic Regression, Naive Bayes, and Random Forest; highest test accuracy achieved was 92%.
- **GridWorld Reinforcement Learning with SARSA and Q-Learning Agent Policies** (Github Link): Implemented SARSA and Q-Learning agent policies in a Python environment for reinforcement learning within a GridWorld simulation. Designed agents to navigate a 4x4 grid, optimizing strategies to avoid pitfalls and reach the goal, utilizing SARSA for on-policy learning and Q-Learning for off-policy learning.
- **Deep Neural Networks from Scratch** (Github Link): Created MLP and CNN elements from scratch (only using NumPy), enabling forward and backward passes across layers (Fully Connected, Conv2D, GeLU activation, MaxPool, Dropout). Applied SGD and ADAM with cross-entropy loss, weight decay, alongside L1 & L2 regularization for optimal model training.
- **Time Series Forecasting of Crude Oil Prices** (Github Link): Performed time series forecasting in R to predict crude oil prices in the next 10 years using ARIMA, SARIMA, ETS, STLF, and Holt-Winters models. Evaluated model adequacy using MAE and utilized libraries like zoo, forecast, ggplot2, and tseries.

CERTIFICATIONS

- AWS Certified Cloud Practitioner (CLF-02)
- AWS Certified Solutions Architect – Professional Certification
- Deep Learning Specialization (Coursera)