Ram Charan Polisetti

New York, United States | ramchara@buffalo.edu | (716) 256-7258 | www.linkedin.com/in/ramcharanpolisetti | https://rampolisetti.com/

EDUCATION

MS in Data Science | University at Buffalo

December 2023, Buffalo, New York

Relevant Coursework: Statistical Data Mining, Probability, Algorithm Analysis and Design, Data Analysis, Data Visualization, Machine Learning, Deep Learning, Web Scraping, Applications of Data Science.

SKILLS

Industry Knowledge: Deep Learning, Machine Learning, Data Analysis, Natural Language Processing (NLP), ETL, Data Visualization, Image Processing, Quantitative Analysis, Pattern Recognition, Time-Series Analysis, Financial Modeling, Risk Management, Algorithmic Trading, Market Analysis, Bayesian Statistics, Predictive Modeling.

Programming Languages: Python (Pandas, NumPy, Scikit-learn, Keras, TensorFlow, PyTorch, XGBoost, SciPy, Seaborn, Plotly, Beautiful Soup, Selenium), Distributed Computing (MapReduce), R (Tidyverse, ggplot), SQL, MATLAB, C++, C.

Tools and Technologies: AWS, Docker, Snowflake, Hadoop, NoSQL databases (MongoDB), Tableau, Power BI, Git, Excel, Advanced Financial Analysis Tools.

Interpersonal Skills: Cross-functional collaboration, Stakeholder Management, Adaptability, Process Improvement, and Supply Chain Optimization.

PROJECT EXPERIENCE

Fake News Detection on News Articles

- Conducted an exploratory data analysis of news articles in R, focusing on language patterns, sources, and media bias, and enhanced data quality through comprehensive text preprocessing, enabling more effective natural language processing workflows.
- Designed and tuned a Random Forest machine learning model, incorporating innovative feature engineering with title and body text analysis, which led to an 80% accuracy rate in differentiating authentic from counterfeit news.

Improving Civic Services through Data-Driven Insights: NYC 311 Request Analysis

- Developed machine learning pipeline using SMOTE oversampling combined with RandomUnderSampler for balanced training sets, improving model performance on severely skewed NYC housing complaint data (34:1 ratio). Tuned Random Forest and XGBoost models through stratified cross-validation; achieved over 81% accuracy on test sets.
- Assessed model performance on noisy real-world data employing SMOTE and under-sampling alongside accuracy, AUC-ROC, confusion matrix, and other metrics to select optimal approach for disproportionate class distributions.

Telco Customer Churn Analysis

- Led the comparison and evaluation of six machine learning models, ultimately identifying Gradient Boosting as the most effective for customer churn prediction with the Scikit-Learn library, and enhanced model precision through the engineering of 15 targeted features.
- Improved the Gradient Boosting model's accuracy to 86% by meticulously fine-tuning hyperparameters via RandomizedSearchCV, alongside refining the dataset and utilizing mutual information scores to bolster precision and ensure robustness for future analysis.
- Developed a custom analytics dashboard in Python to visualize churn predictions and insights from Gradient Boosting models, showcasing front-end development skills and the ability to translate complex model outputs into actionable business intelligence.

Comprehensive Analysis of the European Soccer League with SQL

- Analyzed eight years of European soccer data with advanced SQL techniques to reveal a clear correlation between teams' goals scored and their league standings, informing strategic decisions.
- Developed reusable SQL views and custom metrics for detailed season-by-season team performance analysis, and calculated key statistics using SQL functions to support comprehensive evaluations.
- Designed and implemented an SQL-based data retrieval system to automate the extraction of complex performance metrics, demonstrating skills in database schema design and optimization for high query efficiency.

Analyzing and Forecasting Amazon Stock Prices

- Crafted and optimized a SARIMAX forecasting model, achieving a remarkable 0.98 R-squared and a 3.36 RMSE on test data, which marked a significant advancement over existing baseline models for predicting stock closing prices.
- Executed a comprehensive statistical analysis leading to an improved ARIMA model with a 24-point AIC reduction, and conducted Augmented Dickey-Fuller tests, enhancing the SARIMAX model's accuracy with an MSE of 2.46 and revealing a strong 96% correlation between stock opening and closing prices.

Deploying GRU, LSTM, and RNN for Market Analysis

- Pioneered the development of advanced RNN, LSTM, and GRU models for stock price prediction, crafting a predictive feature set based on 60-day historical data that boosted accuracy by 20% and honed model precision to 85%, sharpening the forecast error margin from 20.3% down to 17.2%.
- Mastered GRU model performance, achieving an industry-leading RMSE of 0.0018 after meticulous hyperparameter adjustments, excelling past LSTM and RNN benchmarks, complemented by creating over 20+ dynamic Plotly visualizations to reinforce strategic, data-led trading decisions.
- Implemented a cloud-based deployment of predictive models using AWS services, integrating GRU, LSTM, and RNN architectures into a production environment, emphasizing skills in cloud computing, deployment automation, and distributed computing.

WORK EXPERIENCE

Transportation Specialist | Amazon Development Center

November 2020 - July 2022, Hyderabad, India

- Orchestrated cross-departmental automation enhancements in dispatch systems, boosting scheduling accuracy by 40%, curbing carrier-related fraud by 15%, and recapturing 2.5 daily operational hours.
- Systematized onboarding for transportation operations by authoring SOPs and integrating automation, trimming acclimatization time by 30%, and reinforcing team efficiency.
- Analyzed transportation data using SQL, optimizing route efficiency, which slashed empty miles by 8% and pared down logistical expenses.
- Devised and managed real-time dashboards for logistics oversight, diminishing operational incidents by 20% and enhancing punctuality in deliveries.

CERTIFICATIONS

Google Data Analytics Professional Certificate | Coursera

Intro to Hadoop and MapReduce | Udacity