RAMAN SINGH

+1 928-265-6218 LinkedIn Tableau GitHub singrama@iu.edu

EDUCATION

Indiana University (IU) Bloomington, Indiana, U.S.A.
Master of Science in Data Science Aug 2021 – May 2023

- **Relevant Courses:** Advance Database Concept, Machine Learning for Signal Processing, Data Mining, Information Visualization, Statistical Computing, Reinforcement Learning, Deep Learning Systems, Management Access & Use of Big Data

M.S. Ramaiah Institute of Technology Bengaluru, Karnataka, India
Bachelor of Engineering (B.E.) Aug 2012 – Jul 2016

- **Relevant Courses:** Artificial Intelligence, Mathematics I – IV, Statistical Quality Control, Database and Management System

TECHNICAL SKILLS

Programming Languages: Python, R, SQL, No SQL, C++, VB.Net, Visual Basic for Application (VBA), CAT Script, and VBScript
Python Libraries: TensorFlow, Keras, Librosa, PyTorch, Scikit-Learn, Seaborn, Pillow, OpenCV, Numpy, boto3, Dash, Flash, PySpark
Statistics: Hypothesis Test, A/B Testing, T-Test, p-Test, Type I & Type II Errors, Bayes Theorem, Derivatives and Gradients
Visualization Tools: Tableau, Power BI, Python Plotly, PowerPoint, Gephi, Sci2
Miscellaneous: Docker, Snowflake, Spark, AWS (EC2, Lambda, S3), GitHub, HTML, CSS, UiPath (RPA), Linux, Big data & REST APIs

PROFESSIONAL EXPERIENCE

Paul H. O'Neill School (IU), Bloomington, U.S.A. ***Research Assistant*** Aug 2022 – Present

- To study Transparency in Coverage (TIC) from insurers to reveal prices of healthcare which are hidden from customers before they must make their decisions for various EINs, billing codes and negotiation types and rates
- Converted CMS (Centers for Medicare & Medicaid Services) defined machine-readable file format into CSV file format to understand different negotiated rates by large insurers e.g., United Health Care, Anthem, and Humana for all National Provider Identifiers (NPIs) across various Employers in the state of Indiana for creating visualizations using Tableau
- Using High Performance Computing (HPC), PySpark to expedite the processing of 3-5 gigabytes of nested JSON and CSV files

ANNUITAS, Atlanta, U.S.A. ***Summer and Fall Data Analyst Intern*** Jun 2022 – Present

- Implemented Docker on EC2 (AWS) to automate daily data cleaning from marketing advertisement databases
- Conducted Multi-Interaction Tracking (MIT) audit on data from Adobe Marketo using Natural Language Processing (NLP) models to predict the marketing channel for each marketing record across 5+ customers having different channel criteria
- Architected and migrated 15+ SQL Postgres database to Snowflake to achieve data warehousing and data analytics

Kelly School of Business (IU), Bloomington, U.S.A. ***Research Assistant*** Dec 2021 – Jun 2022

- Predicted the best options market betas(β) to outperform various measures of beta, like in research paper Frazzini and Pedersen (2014) for stocks using features, SVIX (Short Volatility Index), expected returns, dividend strips, implied volatility, strike, and spot price based on research papers Martin (2016), Martin & Wagner (2019), and van Binsbergen (2012)
- Trained a LSTM neural network model using options data from years 1996-2000 to predict market betas for the next days' returns and had 15% lower sum of squared errors with the testing data than competing methods by testing the sample in 2001 – 2020 options data to avoid anomaly returns with optimized beta(β) value
- Used deep learning methods, High Performance Computing (HPC), PySpark, batch processing to implement the LSTM model and reduce computational time by 80% to preprocess, train and test the data

Faurecia India Private Limited, Pune, India ***Knowledge Based Engineer (KBE)*** Jul 2019 – Aug 2021

- Led the development and deployment of 6 machine learning models and data pipelines that saved 9400+ support hours
- Collaborated with executive team and created multiple interactive dashboards using SQL & Power BI to track project Key Performance Indicators (KPIs) e.g., milestones, task list, resource utilization, quality matrix, and ROI

Inteva Products, Bengaluru, India ***Knowledge Based Engineer (KBE)*** Aug 2016 – Jul 2019

- Compiled and analyzed quality reports for 8 cross-functional teams to build a monthly quality dashboard using Tableau and recommend probable improvement areas based on the previous experiences from managers
- Automated monthly dashboard for the Manufacturing Execution System (MES) report of 12 global manufacturing plants using SQL and Tableau which saved 87% of support hours per month and provided insights from the report

MAJOR PROJECTS

Prediction of Judicial Decisions for Human Rights (Scikit-Learn, NumPy, Seaborn, Python, NLTK libraries) Jan 2022 - Mar 2022

- Programmed NLP based predictive model to successfully obtain the outcome of court's cases by comparing several US constitution articles that described the rules of human rights mentioned in the lawsuit
- Obtained 79% accuracy by implementing bag of N-grams to train the Support Vector Machine (SVM) classifier coupled with linear kernel, regularization parameter, 10-fold cross validation and grid search

Amazon Product Recommendation System (PyTorch, NumPy, Batch processing, HPC, Seaborn, Python) Aug 2022 – Present

- Used a novel approach to build a content-based recommendation system with cognitive filtering and additional features such as ratings and user comments for all the books present in the Amazon database from 2018 to 2022
- Developed gated graphical neural network (GGNN) on user comments over a layer of natural language processing to suggest new products from the amazon website. Used A/B testing and cosine similarity as the evaluation metric