Karthikeyan Rajagopal

New York, United States | +17169079288 | rajagop3@buffalo.edu | LinkedIn | GitHub | Website

PROFESSIONAL SUMMARY

Data Scientist with 4 years of experience having a strong background in optimizing context-specific product recommendations, utilizing Machine Learning methods and Predictive analytics to solve complex business and engineering challenges and drive growth in product development, enhancing customer and product metrics and campaign performance reporting using data analytics and dashboards. Ideated data-driven solutions to improve customer acquisition, customer engagement, conversion, and retention.

SKILLS & LEADERSHIP

Programming Language: Python, Java, SQL, Scala, C, C++, R, Golang
Technology: Git, AWS, Azure, GCP, JIRA, Excel, Spark, Hadoop, Apache Airflow, Hive, Kafka, Power BI, Tableau
Libraries: Pandas, Numpy, scikit-learn(sklearn), Matplotlib, scipy, TensorFlow, OpenCV, PySpark, PyTorch, Spacy, Transformers
Technical Skills: Machine Learning, Natural Language Processing (NLP), Recommender Systems, Exploratory Data Analysis (EDA), Data Mining, Analytics, Data visualization, Statistics, Clustering, Forecasting, A/B testing, Deep Learning, Dashboard, ETL
Soft Skills: Communication skills, Analytical skills, Project Management, Innovative, Results-oriented, Leadership, Decision making

PROFESSIONAL EXPERIENCE

Graduate Research Assistant| SUNY Research Foundation | Buffalo, New York **APRIL 2023 – PRESENT**

Optimization of Microservices Architecture for Domain-Specific Recommendations: *Python, Golang*

- Refactored Golang microservices to Python, leading to enhanced architectural integrity and improved maintainability
- Utilized NLP techniques with Large Language Models (LLMs) to refine semantic similarity, resulting 32% surge in data extraction
- Engineered IDF based scoring algorithm to enrich domain-specific scientific research recommendations by 17%

Data Scientist | Crayon Data | Chennai, India **JUNE 2018 – JAN 2021**

Built client specific applications to increase productivity and customer retention: *Python, Scala, SQL, Power BI*

- Remodeled company's patented algorithm to refine personalized recommendations increasing customer coverage by 300%
- Devised a robust nearest neighbor algorithm personalizing customer segmentation and boosting targeted marketing by 24%
- Optimized delivery routes increasing $2M in revenue and designed an interactive Power BI Dashboard to track essential KPIs

Refurbished an end-to-end ETL and Machine Learning pipeline: *Python, Scala, Java, SQL, Spark*

- Led a team of 5 to modularize full-fledged big data pipeline to enable client-specific and context specific customization
- Designed application with hyper-parameters to fine tune each module and thereby reducing resources and time used by 100%
- Streamlined data preprocessing with advanced entity resolution techniques increasing accuracy of predictions by 15%

Improved Travel Recommendation of Middle East's largest global Airlines: *Python, Scala, SQL, R, Tableau*

- Derived key metrics with 10+ stakeholders cross-functionally, and sprint planning to improve end-to-end customer experiences
- Performed sentiment analysis to extract and score relevant tags from travel data to construct centralized repository
- Developed user-based and item-based collaborative filtering using matrix factorization and querying to enhance metrics by 37%

ACADEMIC PROJECTS

Continuous Checkpointing for Deep Learning Recommender Systems | University at Buffalo **JAN 2023 – MAY 2023**

- Collaborated with META to develop novel checkpoint mechanism for fault tolerance and seamless recovery from system failure
- Implemented distributed architecture in AWS EC2 instance to scale for production deployment and handle terabytes of dataset
- Leveraged Kafka and PyTorch Hooks to optimize extraction of weights during backpropagation, reducing save time by 42%

Evaluating Multiple Machine Learning Classification Algorithms | University at Buffalo **JUL 2022 – DEC 2022**

- Integrated PCA and LDA methodologies to preprocess and transform high-dimensional data into a lower-dimensional space
- Benchmarked performance of perceptron, decision tree, and neural network computing F1 score, AUC ROC and cross validation
- Fine-tuned model hyperparameters with L2 regularization and dropout to enhance model performance to 81%

Injury Severity Among Pedestrians: A Data Mining Approach | IIT Madras | Master's Thesis **AUG 2017 – MAY 2018**

- Analyzed various accidents based on different inputs to detect severity of accidents employing statistical methods
- Identified major risk factors by performing feature engineering with empirical analysis and extensive logistic regression
- Achieved an AUC of 0.77 in ROC curve for statistical modeling of severity risk in pedestrian motor accidents

EDUCATION

University at Buffalo - State University Of New York | MS in Computer Science (GPA 3.79/4.00) **|** Buffalo, USA **AUG 2022 – DEC 2023**
University of Melbourne | Graduate Diploma in Data Science **|** Melbourne, AUSTRALIA **MAR 2021 – DEC 2021**
Indian Institute of Technology Madras | B. Tech & M. Tech (Automotive Engineering) **|** Chennai, INDIA **JUL 2017 – MAY 2018**