

# HARI PRASAD RENGANATHAN

1520 Riverside Station Blvd, Secaucus, NJ

[linkedin.com/in/hariprasad20/](https://www.linkedin.com/in/hariprasad20/)

917-238-4627

[hr2514@columbia.edu](mailto:hr2514@columbia.edu)

## EDUCATION

### Columbia University

New York, NY

#### Master of Science, Data Science

Dec 2022

Coursework: Machine Learning for Data Science, Exploratory Data Analytics and Visualization, Algorithms for Data Science, AI Applications in Finance, Finance and Structuring for Data Science

### Kumaraguru College of Technology

Coimbatore, India

#### Bachelor of Engineering, Mechanical Engineering

Jun 2021

*Awarded with meritorious scholarship (\$600) 3 times for excelling in top 5% of the department*

## SKILLS

- **Tools:** Python, R, MS Excel, Tableau, SQL, WordPress, Snowflake, AWS, PySpark, Databricks
- **Libraries/Frameworks:** Pandas, Numpy, Matplotlib, Seaborn, Scikit-learn, Keras, Beautiful Soup, ggplot2, Git, Dash
- **Data techniques:** Neural Networks, Supervised and Unsupervised Machine Learning, Exploratory Data Analysis, Web Scraping, Data Visualization, Hypothesis Testing, Time Series, ETL, Sales forecasting, Price Elasticity, Fuzzy Matching

## EXPERIENCE

### Take-two Interactive Software, Inc.

New York, NY

#### Associate Data Scientist

Aug 2022 - Present

- Redesigned model store client, enabling version control of model metadata and implementing restrictions on direct model registration into development, enhancing model governance and improving collaboration between teams.
- Developed an application to compare 1400 machine learning models simultaneously and implemented batch registration of models, streamlining workflow efficiency. Reduced model deployment time by 80%.
- Executed simulations of sales for various discounts, providing critical insights for promotion planning and enabling the sales team to make informed decisions, resulting in a projected increase in sales revenue by 15%.
- Created a fuzzy matching algorithm utilizing cosine similarity to join tables with external vendor data, improving data accuracy and enabling more informed decision-making. Increased data matching accuracy by 95%.
- Implemented a robust model validation pipeline, allowing for the tracking and evaluation of model performances at the entity level (~5000 models), leading to continuous model improvement and enhanced business outcomes.

#### Data Science intern

May 2022 - Aug 2022

- Optimized sales forecasting model by introducing ARIMA, enabling more accurate predictions for sales projections. Increased revenue forecast accuracy by 25%.
- Collaborated closely with ML Engineers and Data Analysts to deploy price elasticity model outputs into production using Dash, resulting in real-time insights for the sales team and enhancing decision-making capabilities.
- Conducted comprehensive competitive analysis on data from 72 companies using Snowflake, SQL, and Tableau, providing valuable insights for identifying market trends, opportunities, and potential risks.

#### Ask2AI

New York, NY

#### Financial Data Scientist Spring Intern

Jan 2022 - May 2022

- Forecasted future returns and volatility using a comprehensive 650-dimensional macro-economic time-series dataset. Achieved an average forecasting accuracy of 85% for returns and 80% for volatility.
- Applied rolling window Principal Component Analysis (PCA) to identify the most influential factors driving the macro-economic trends. Reduced dimensionality of the data by 90% while retaining 95% of the explained variance.
- Built a predictive model using LSTM Neural Networks to capture complex temporal dependencies in the data. Achieved an impressive Mean Absolute Percentage Error (MAPE) of 7% in predicting future returns and volatility.

#### GoTTR

New York, NY

#### Founding member & Data Scientist

Jul 2021 - Dec 2021

- Regularized data-scraping and data-modeling for a massive 100 GB dataset, reducing manual work that previously took 4 hours to just a few seconds, resulting in significant time savings and improved productivity.
- Led a team of 5 in computerizing data management on Google Cloud (GCP) and implementing an Email-Alert system. This initiative resulted in monthly cost savings of \$7000 by improving operational efficiency.
- Leveraged Python, SQL, Pandas, and Scikit-learn to automate data management processes, achieving a 75% reduction in data processing time and improving data quality and accuracy.

#### DADDYODA

Coimbatore, India

#### Co-founder & Marketing Data Scientist

Jul 2020 - Jul 2021

- Automated the delivery of 50+ weekly promotional emails to a user base of 10,000+, utilizing Python. Increased operational efficiency by 30% and reduced manual effort by 20 hours per week.
- Implemented text analytics to predict customer sentiments in reviews, reducing Customer Acquisition Cost (CAC) by 20% through improved targeting and personalized marketing strategies.