

DIVYANK RAHORIA

Los Angeles, CA | divyank.d.rahoria@gmail.com | +1 (858) 336-1044 | [LinkedIn](#)

Summary

- 4+ years of experience in Data Science, Machine Learning, and Statistical Analysis. Proficient in Python, SQL, ML, and Deep Learning algorithms like Regression, Clustering, Decision Trees, Random Forests, Time Series, CNN, and RNN.

Experience

Data Scientist | Newegg Inc, California

Jun 2021 - Current

- Executed test-and-learn strategies and A/B testing to evaluate new product features. Analyzed results and presented recommendations to cross-functional teams.
- Built models using Python and SQL to uncover customer insights from transactional data. Identified opportunities to optimize customer segmentation and marketing.
- Led analytics projects analyzing large datasets and developing inventory forecasting models. Communicated data-driven insights through presentations.
- Established data pipelines and architecture. Used BigQuery and other big data tools for processing and analysis.
- Built machine learning models targeting key business issues that balanced predictive power and ease of maintenance/operations.
- Mentored junior data scientists in enhancing their statistical modeling and data analysis skills to elevate team effectiveness.

Data Scientist | Remote (Part-Time) | Lavit Pure Agro, India

Jan 2021 - May 2021

- Created a robust anomaly detection system to monitor the company's supply chain data, identifying unusual patterns such as unexpected delays in shipments, sudden changes in costs, or irregularities in supplier performance.
- Developed and implemented predictive models using machine learning algorithms such as linear regression, classification, multivariate regression, and Random Forests.

Graduate Researcher | ECE Department, UCSD

Mar 2020 - Jun 2020

- Preprocessed the American sign language data and compared different deep learning models like CNN, VGG16, Inception Net, RESNET for the task of predicting ASL translation labels on test data. [Paper Link](#)

Graduate Researcher | ECE Department, UCSD

Mar 2019 - Jun 2019

- Improved the accuracy of VQA using never before used deep learning model based on RNNs and incorporated GRU to boost the accuracy over previous work by significant amount while using CUDA for faster training.

Data Analyst | PWD, India

July 2017 - Aug 2018

- Worked on Excel, automated multiple reports and cleaned them using Python scripts by working on the combination of unstructured and structured data from multiple sources and visualized using Tableau.

Research Intern | Bangkok University Robotics Lab, Thailand

Jan 2017 - Jun 2017

- Created Python scripts to develop portable navigation system device for visually impaired people (Thailand Association of Blind).

Education

University of California San Diego (UCSD)

2020

M.S. in Computer Science (Major in Artificial Intelligence)

VIT University

2017

B. Tech, Electronics and Communication Engineering (Computer Engineering)

Projects

Gesture Genie

Mar 2020 - Jun 2020

- Incorporated the Piano Genie by Magenta and Google's Hand Tracker to develop [Gesture Genie](#) by working on Javascript based Magenta API to aid in deep learning based music research.

Airbnb Price & Superhost Prediction

Jan 2019 - Mar 2019

- Analyzed exploratory data for Airbnb Paris dataset to predict price and superhost qualities prediction using different regression models including Linear Regression, Ridge Regression, Gradient Boosting, Random Forest Regression to get 91 % accuracy on test data.

Text Classification and Sentiment Analyzer

Mar 2019 - Jun 2019

- Classified the user input in 5 different categories based the predicted topic related to the input and performed sentiment analysis on it.
- Found top k words responsible for positive/negative impacts and explained all 5 categories using Lime architecture.

Skills

Python, R, Java; Gen AI, LLMs, Regression, Clustering, Random Forest, NLP, Decision Tree, Time Series, Sequences, A/B Testing, CNN, RNN; Big Query, SQL Server, MySQL, Hadoop, Spark; Vertex AI, Big Query, Git, CUDA, Keras, PyTorch, Tableau, Power BI.