

Viraj Thakkar

☎ (+1) 646-409-9883 | ✉ viraj.thakkar@nyu.edu | 🏠 virajthakkar.com | 📷 virajthakkar | 📺 virajthakkar285

Education

New York University

New York, NY

MS IN DATA SCIENCE | GPA: 3.7/4

Aug 2019- May 2021

- Relevant coursework: Machine Learning, Big Data, Probability and Statistics, Linear Algebra, Deep Learning, NLP.
- Graduate Assistant and Grader at NYU Courant Institute of Mathematical Sciences and NYU Center for Data Science.

Homi Bhabha National Institute

Mumbai, India

INTEGRATED MASTER'S DEGREE(BSc+MSc) | MAJOR: PHYSICS, MINOR: MATHEMATICS | GPA: 7.89/10

Aug 2014 - May 2019

- Relevant coursework: Mathematical Methods I and II, Stochastic Processes, Differential Equations, Numerical Analysis, Quantum Mechanics, Statistical Mechanics, Einstein's Theory of Relativity, Non-linear physics and Chaos.

Skills

Languages Python, C++, MATLAB

Tools and Technologies NumPy, pandas, scikit-learn, Matplotlib, keras, SQL, Statsmodels, Tableau, spaCy, Hadoop, Spark, Git, LaTeX

Machine Learning Artificial Neural Network, Decision Trees, K-means clustering, Regression analysis, Generalized Linear Models

Statistics Neyman-Pearson theory of hypothesis testing, Bayesian Statistics, Stochastic processes

Experience

Tenaska, Inc.

Dallas, TX

DATA SCIENTIST

Febr 2022 - Present

- Critical contribution in the development and improvement of real-time optimization and bidding strategies of batteries in the electricity market of Texas and California. Resulted in an order of magnitude of decreased deviations, improvement in the reliability of the batteries, and increased revenue.
- Development and maintenance of critical benchmarking metrics and performances. Conducting extensive testing under various configurations and parameters to evaluate new features and propose enhancements, crucially impacting the performance and driving daily operational decisions.
- Predicting natural gas prices and probabilistic time-series forecasting of real-time nodal electricity prices using deep learning models.

HealthBadge, Inc.

Remote

DATA SCIENTIST INTERN

Aug 2021 - Jan 2022

- Working on building a GUI to obtain disease statistics from multiple sources for an Epidemic Simulation Framework using web scrapping and NLP.
- Using Named Entity Recognition and Outlier detection to get accurate disease statistics from various trusted sources.

RUMO S.A.

New York, NY

DATA SCIENTIST INTERN

Sept 2020 - Dec 2020

- Developed a method to improve fuel consumption efficiency of the RUMO railway line operations in Brazil which can lead upto \$200k savings p.a.
- Designed a methodology to find malfunctioning locomotives when the amount of diesel refueled is more than estimated consumption.
- Achieved a recall of 70% for detecting unusual refueling events where fuel consumption of the locomotive is more than the estimated consumption.

Homi Bhabha National Institute

Mumbai, India

GRADUATE RESEARCHER

May 2018 - May 2019

- Worked with the dark matter search experiment-SuperCDMS collaboration with the aim to distinguish nuclear recoils(signal) and electron recoils(background) events in the data by using Machine Learning classification.
- Extracted and reconstructed features from the data by remotely accessing to SLAC National Accelerator Laboratory computing resources.
- Improved signal to background ratio by 5.3% by combining Decision Trees algorithm with traditional background filtering methods. [View link](#)
- Implemented Supervised Machine Learning classification of signal and background particles in Large Hadron Collider(LHC), CERN experiment.
- Reduced background particles in the spectrum by using adaboost and improved the signal to background ratio by 25.55%. [View link](#)

Projects

Predicting Option prices using Deep Learning Networks

- Simulated stock prices using geometric Brownian motion process.
- Calculated option prices for various combinations of risk-free rate, volatility, maturity and strike prices under the Black-Scholes model assumptions.
- Predicted option prices using Deep Learning network with a root mean square error of 2 cents per dollar of strike price. [View link](#)

Intraday currency trend size prediction using Machine Learning

- Collaborated in predicting larger than average foreign exchange currency rate fluctuation for the USD-Mexican Peso pair by using machine learning.
- Selected the best model using time series cross-validation method and calculated the probability of observing a larger than average trend size of currency exchange rate given the current market conditions.
- Implemented decision trees with boosting as the classifier which resulted in the best performance with an AUC of 0.616. [View link](#)