

# VIBHAV MISRA

New York, NY 11206 | [vibhavemail@gmail.com](mailto:vibhavemail@gmail.com) | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

## EDUCATION

**Pace University, Seidenberg School of Computer Science and Information Systems**

**New York City, NY**

Master of Science (MS) in Data Science | **Concentration:** Data Science and Analytics

May 2026

**Chandigarh University, Apex Institute of Technology**

**India**

Bachelor of Engineering (BE) in Computer Science with a specialization in Artificial Intelligence and Machine Learning

June 2024

## PROFESSIONAL EXPERIENCE

**EdMyst Inc.**

**Lewes, DE**

**Data Science Intern**

**June 2025 - September 2025**

- Designed and documented a personalization system for an AI coaching platform (Edy), integrating onboarding data, session performance, behavioral metrics, and feedback into a long-term user memory, with supporting architecture and data flows using MongoDB, AWS S3, Lambda, and LLM prompt injection.
- Developed and tested a real-time video chunking and upload pipeline using WebRTC, Node.js, and AWS S3 multipart uploads, enabling seamless video transfer and future analytics integration.

## ACADEMIC PROJECTS

**Exoplanet Habitability Explorer**

**September 2025**

- Built an interactive Streamlit app using NASA's Exoplanet Archive to explore 5K+ exoplanets with a transparent habitability score and tunable weight-presets.
- Engineered astrophysical features (luminosity proxy, semi-major axis, estimated insolation) and trained a Random-Forest classifier to flag "optimistic habitable-zone" candidates; integrated ML predictions and metrics (ROC-AUC  $\approx$  0.8) into the UI.

**Will I Catch That Train? - Real-Time NYC Subway Tracker**

**August 2025**

- Built a Streamlit web app that predicts ideal departure time for NYC J/Z/M subway riders by combining MTA GTFS-RT live arrivals with static GTFS schedules.
- Developed a time-based interpolation algorithm to simulate real-time train positions on a map, displaying colored route-specific icons and arrival boards with actionable "Leave-Now" recommendations.

## SIMULATED INDUSTRY PROJECT

**British Airways Data Science Job Simulation**

**Forage - November 2024**

- Scraped and analyzed over 1,000 customer reviews using sentiment analysis techniques to uncover key insights, with 62.2% of reviews being positive and common themes centered around service, flight delays, and amenities.
- Built a predictive model using Random Forest to identify key factors influencing booking completion, achieving an accuracy of 85%.

## TECHNICAL SKILLS

**Programming:** Python, SQL, R, React, JavaScript

**MLOps & Experiment Tracking:** MLflow, Model Registry, Experiment Tracking

**Cloud Platforms:** AWS (S3, Lambda, EMR on EC2)

**ML & AI:** Supervised & Unsupervised Learning, Neural Networks, Time Series Forecasting, Generative AI (OpenAI API)

**Data Science & Analytics Tools:** Pandas, NumPy, Scikit-Learn, TensorFlow, Prophet, Keras, Jupyter Notebooks, Anaconda

**Big Data & Databases:** MySQL, MongoDB, HBase, Neo4j, Hadoop (HDFS, MapReduce), Apache Spark

**Data Visualization:** Power BI, Looker (dashboards, basic LookML modeling), Tableau, Matplotlib, Seaborn, Excel

## CERTIFICATIONS

- Mathematics for Machine Learning and Data Science Specialization** (Linear Algebra for Machine Learning and Data Science, Calculus for Machine Learning and Data Science, Probability & Statistics for Machine Learning & Data Science) **Coursera - October 2023**
- Natural Language Processing Specialization** (NLP with Classification and Vector Spaces, NLP with Probabilistic Models, NLP with Sequence Models, NLP with Attention Models) **Coursera - May 2023**
- Applied Data Science Specialization** (Python for Data Science, AI & Development, Python Project for Data Science, Applied Data Science Capstone, Data Visualization with Python, Data Analysis with Python) **Coursera - July 2022**

## ACADEMIC ACHIEVEMENTS

- Part of the pilot batch for the online BSc Degree in Programming and Data Science offered by Indian Institute of Technology Madras.
- Completed the 8 Foundational Level Courses (32 credits) and earned a Foundational Certificate from CODE, IIT Madras.
- Progressed through Diploma Level courses - Including: Machine Learning Foundations, Business Data Management, Machine Learning Techniques, Business Analytics and Tools in Data Science. Transitioned to Master's program in the USA before completing the Diploma Level.

## LANGUAGES

English (Fluent), Hindi (Native)