

VARSHA VATTIKONDA

+1 (805)-743-9755 | varshavattikonda@gmail.com | [LinkedIn](#)

SKILLS

Languages	Python, SQL, C#, R
ML Frameworks & MLOps	PyTorch, Tensorflow, Scikit-learn, nltk, AWS SageMaker, GCP (Vertex AI), Airflow
Data Processing	Azure (Data Factory), Spark, Databricks, Pandas
Tools	MSSQL, Splunk, Power BI, Salesforce

WORK EXPERIENCE

Machine Learning Engineer - JP Morgan Chase, New York City, USA

Jun'23 - Curr

- Drove a **\$5M** uplift in annual revenue by partnering with Sales and engineering an outreach system; **Random Forest** model to identify overlooked contacts and deliver leads via a REST API, boosting sign-ups by **5%**
- Architected a scalable **MLOps** framework for **50+** data scientists, creating reusable components for training, model versioning, and automated deployment, reducing model-to-production lifecycle by **50%**
- Developed a recommendation engine uncovering **30%** more qualified leads for follow-on offerings; an **XGBoost** trained on 13F, trading, and CRM data, deployed as a real-time prediction service.
- Engineered a global alerting platform for **300+** sales reps that used a **3-layer MLP** (Neural Network) trained on **AWS SageMaker** to score and prioritize alerts, increasing user engagement by **25%**
- Productionised a time-series stock recommendation engine using an **LSTM+Attention** model, architecting a batch prediction pipeline processing **10GB+** of trade history and serving with **80% Average Precision**
- Designed and deployed a client meeting allocation service using a constrained optimization algorithm, which was served via a REST API, improving meeting fulfillment by **18%**.

Machine Learning Engineer - MSCI Inc., Mumbai, India

Aug'16 - Jul'21

- Generated **\$1.2M** in ARR by developing a batch recommendation system using **PCA and K-Means** to cluster clients and generate cross-sell opportunities via REST API integrated with CRM
- Reduced customer churn by **10%** by productionizing an **XGBoost** prediction model; architecting a feature engineering pipeline in SQL to compute behavioral trends and deploying the model as a batch scoring job
- Deployed a two-stage prediction model (**K-Means + Linear Regression**) as a real-time REST API to estimate simulation durations, providing accurate (**MAE of 5 minutes**) time-to-completion estimates
- Architected a self-healing ETL framework to ingest **10GB/day** from 10+ sources, providing clean, feature-ready data for all production ML models, reducing data prep time by **15%**
- Engineered a sentiment analysis pipeline that processed 10,000+ client emails using an **LSTM** model deployed on **Azure Functions**, providing real-time sentiment signals

RESEARCH & INTERNSHIPS

Research Assistant - New York University, USA

Jan'23 - May'23

- Architected and built a multi-stage **NLP pipeline** to automate the extraction of causal relationships, orchestrating **BERT**-based models for sequence and token classification
- Engineered a data annotation tool with **JavaScript** in Qualtrics to streamline the creation of a high-quality training dataset, reducing the data labeling lifecycle by **40%**

Data Science Intern - MSCI Inc. ESG NYC, USA

May'22 - Dec'22

- Engineered and deployed an end-to-end **Geospatial ML system** on **Google Vertex AI**, which served as the foundational PoC for a **\$1M** strategic partnership between MSCI and Google
- Built the production pipeline to perform scalable batch-inference on **300,000+assets**; data ingestion using **Google Earth Engine API** and a two-stage ML model (**KMeans Clustering + Random Forest**)

EDUCATION

M.S. Data Science | New York University, USA (3.9/4)

Sep'21 - May'23

Coursework : Natural Language Processing, Deep Learning (Computer Vision), Big Data (Spark)

BTech & MTech Electrical Engineering | Indian Institute Of Technology Madras (IITM)

Aug'11 - May'16

Coursework : Data Structures & Algorithms, Machine Learning, Linear Algebra, Probability and Statistics