+1(240)-415-2871 | siddhant2805@gmail.com | github.com/Siddhant2805 | linkedin.com/in/siddhant-chahana-b202228283/

Education

University of Maryland, College Park May. 2025
Master of Science in Applied Machine Learning *GPA: 3.65/4.0*

Vellore Institute of Technology, Vellore April 2019
Bachelor of Technology, Mechanical Engineering *GPA: 3.5/4.0*

About Me

Machine Learning Engineer with 4 years of industry experience and a graduate degree in Applied ML. Skilled in building robust ML pipelines for both GenAI (RAG-based code generation and summarization) and traditional ML (RNN forecasting in biomechanics). Experience in deploying churn prediction and recommendation models on AWS SageMaker.

Experience

Dynamics and Control Lab, UMD College Park *Machine Learning Engineer* Feb. 2024 – Present

- Developed forecasting models using RNNs for automating a biomedical device used for rehabilitation in stroke patients.
- Multi-time step prediction for weight and torque in drilling processes using Higher Order SVD and Gaussian Process Regression.
- Preprocessed digital signal data to analyze gait patterns and cluster them using k-nearest neighbor algorithm.
- Generated synthetic time series data using GANS for dataset augmentation.

Godfrey Phillips India Ltd. *Data Scientist* Aug. 2021 – Apr. 2023

- Multivariate Time Series forecasting using ARIMA and SARIMA. Time series classification and clustering done for brand-wise material procurement planning.
- Sales analytics and region clustering using k-means and DBSCAN for optimizing merchandising.
- Deployed ETL pipeline for extracting SQL queried data and created sequential tensors for forecasting using LSTM networks.

Godfrey Phillips India Ltd. *Operations Analytics* Jan. 2020 – Aug. 2021

- Production Planning and control, supply chain and operation analytics and dashboard creation for procurement and inventory management.

Projects

Image Captioning Model *Seq2Seq, Transformers, PyTorch, Keras, CUDA*

- Developed an image captioning model utilizing Transformer-based architecture to generate descriptive captions for images. Used patch-based processing to enhance context learning.

GitSummarizer - RAG *RAG, Huggingface, Ollama, Pinecone, AST, Python*

- Built a Retrieval-Augmented Generation system to enable natural language querying over GitHub codebases using OpenAI embeddings and AST-based parsing. Integrated Pinecone for fast vector retrieval.

Drill Data Forecasting *HOSVD, GPR, Tensor Algebra, MATLAB, Python*

- Created 3D tensors from spatio-temporal drilling data, reduced dimensionality using Higher Order SVD, and performed multi-step forecasting with Gaussian Process Regression.

Skills

Languages & Version Control: Python, R, C++, MATLAB, Linux CLI, Git
Cloud & ML Platforms: AWS SageMaker, Snowflake, Azure, GCP
Libraries/Frameworks: scikit-learn, TensorFlow, PyTorch, Keras, Pandas, NumPy, FastAPI, MLFlow, Docker, Kubernetes, SQL, PostgreSQL, MongoDB
Specialized: GenAI (RAG), NLP, CV, Time-Series, MLOps, LangChain, Tableau, Power BI