

Sidhant Chanana

+1(240)-413-2871 sidchan2805@gmail.com github.com/Sidchan2805 linkedin.com/in/sidhant-chanana-52022b283/

EDUCATION

University of Maryland, College-Park

Master of Science in Applied Machine Learning

Expected May 2025

GPA: 3.8/4.0

Vellore Institute of Technology, Vellore

Bachelor of Technology, Mechanical Engineering

April 2019

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 to enhance material selection 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 create sequential tensors for forecasting using sequence neural networks like LSTMs.

PROJECTS

Image Captioning model | *Sequence Neural Networks, Transformer models, Attention models, PyTorch, Keras, CUDA*

- Developed an image captioning model utilizing a Transformer-based architecture, to generate descriptive captions for images. Processed visual data by dividing images into patches, enabling the model to effectively learn and generate corresponding textual descriptions.

GitSummarizer- RAG | *RAG, Text Embedding, Ollama, Huggingface, LLMs, Python, Abstract Syntax Trees, C++, Git*

- Engineered a Retrieval-Augmented Generation (RAG) system leveraging OpenAI embeddings and LLMs to enable natural language querying of GitHub codebases, achieving context-aware code understanding through parse tree analysis. Used Abstract Syntax tree for object aware code parsing and tokenization and Pinecone for scalable vector search solution.

Drill Data Forecasting | *Higher Order SVD, Gaussian Regression, Tensor, Numpy, Data Assimilation, Dimensionality Reduction*

- 3-d tensor creation from spatio-temporal data of the mechanical process using Numpy and Pandas. Also used MATLAB for coding tensor flattening and algebra functions.
- Higher order dimensionality reduction of data using tensor algebra. Used Gaussian process regression for multi-step ahead time series forecasting.

SKILLS

Languages and 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, Seaborn, FastAPI, MLFlow, Docker, Kubernetes, SQL, PostgreSQL, MongoDB, NoSQL, MySQL

Specialized Skills: GenAI (RAG), NLP, Computer Vision, Time-series analysis, Statistical Analysis, A/B Testing, Market Value Algorithms, MLOps, Power BI, Tableau