

SANTOSH REDDY

ML ENGINEER | DATA SCIENTIST

 +91 765 898 7658  Hyderabad, Telangana  santoshr3dd4@gmail.com

 linkedin.com/in/santoshml  github.com/AkaDeipu

PROFILE

Passionate Machine Learning Engineer with 4+ years of hands-on analytics expertise and proven success in architecting end-to-end ML pipelines—from NLP-driven sentiment analysis to hyperparameter-tuned models deployed via Streamlit for real-time impact. Excel in Python, scikit-learn, TensorFlow, Keras, and MLOps workflows to convert raw data into scalable, high-accuracy solutions that drive business growth. Eager to innovate at the intersection of AI and strategy in fast-paced, high-stakes environments.

SKILLS

Programming & ML: Python, Pandas, NumPy, Scikit-learn, Seaborn, Matplotlib, TensorFlow, Keras, Transformers, Machine Learning Algorithms, XGBoost, LLM, LSTM

Cloud & Development Tools: Azure ML Studio, Google Colab, Jupyter Notebook, VS Code

Data Tools: SQL, Power BI, Advance Excel (Power Query, Power Pivot)

Soft Skills: Problem-Solving, Critical Thinking, Collaboration, Attention to Detail, Ad Hoc Analysis

EXPERIENCE

Ai variant | Sep 25 - Present

Machine Learning Engineer Intern

- Built ML solutions across multiple datasets using core models (Regression, Tree-based, SVM, KNN, Naive Bayes) and advanced techniques (XGBoost, LightGBM, ensemble methods).
- Developed an end-to-end NLP pipeline with multilingual translation, text preprocessing, feature engineering, and TF-IDF & GloVe embeddings for sentiment analysis.
- Implemented time-series forecasting using baseline models and GRU, finalizing an LSTM model through sequence engineering (lookback, horizon) achieving RMSE ≈ 2.
- Trained deep learning models with TensorFlow/Keras, performing hyperparameter tuning (layers, units, batch size, epochs, learning rate) to improve model stability.
- Deployed models via Streamlit (NLP) and Dockerized FastAPI (forecasting), using GitHub and Render, following MLOps best practices.

Vijetha Supermarkets | Mar 21 - Sep 25

Data Analyst

- Built and optimized Power BI dashboards, improving operational efficiency by 30% through real-time KPI visibility and optimized last-mile operations, reducing delivery times by 50%.
- Conducted SQL and Excel-based ad hoc analyses that reduced customer churn by 15%.
- Automated ETL workflows using Power Query, cutting data processing time by 30%.
- Partnered with cross-functional teams to define KPIs and deliver insights that improved decision-making.
- Developed customer segmentation models that increased marketing effectiveness by 30% and boosted acquisition by 10%. Produced customer support analytics that reduced resolution time by 30%.
- Delivered insights driving new product features, resulting in a 25% increase in user engagement.

PROJECTS

Sentiment Analysis of Multilingual Product Reviews (NLP | ML | DL | MLOps)

- Processed 2K+ multilingual customer reviews by performing language detection, translation (GoogleTranslator), and Hindi transliteration.
- Applied advanced text cleaning, lemmatization, stop-word handling, and dual preprocessing pipelines (spaCy + NLTK).
- Engineered features using TF-IDF (unigrams & bigrams), followed by KMeans clustering and dimensional analysis.
- Built and fine-tuned classification models including Logistic Regression, SVM, Random Forest, KNN, Naive Bayes, and ensemble approaches.
- Implemented SMOTE to address class imbalance improving balanced accuracy.
- Developed and optimized neural networks using Keras with dropout regularization and hyperparameter tuning.
- Integrated Transformer-based model (RoBERTa) for high-accuracy multi-class sentiment detection.
- Generated custom word embeddings using GloVe (100-dim vectors) for semantic vector modeling.
- Deployed the final model using Streamlit via GitHub.

Apple Stock Price Forecasting (Time Series | Deep Learning | MLOps)

- Performed end-to-end time series forecasting on Apple stock prices, starting with classical models (ARIMA, GRU) and finalizing an LSTM-based model, which achieved a strong RMSE of 2, outperforming baseline approaches.
- Conducted data exploration, preprocessing, scaling (MinMaxScaler), and windowing, converting historical price data into supervised learning sequences using a configurable lookback and forecast horizon.
- Designed and trained an LSTM neural network using TensorFlow/Keras, handling temporal dependencies and non-linearity inherent in financial time series data.
- Addressed key time-series challenges such as non-stationarity, sequence length selection, and error accumulation by iterative forecasting and careful horizon design.
- Wrapped the trained LSTM model into a custom scikit-learn-compatible estimator, enabling pipeline-based usage and reproducibility.
- Built a production-ready ML pipeline using sklearn pipeline, serialized with Joblib, and exposed predictions via a FastAPI REST endpoint for real-time forecasting .
- Containerized and standardized the environment using Docker and requirements management, demonstrating applied MLOps practices from model training to deployment-ready inference.

EDUCATION

2018 - 2019 **Master of Business Administration**
Cardiff Metropolitan
University

2014 - 2017 **Bachelor of Commerce (Computers)**
Osmania University

CERTIFICATION

- ExcelR — Data Science with Gen Ai (2025)**