SHANMUKHA SARAT PONUGUPATI

910-228-8701|Boston, MA | shanmukhasaratp@gmail.com |Github |LinkedIn

EDUCATION

Northeastern University, Boston, MA **May 2023**

Master of Science in Data Analytics Engineering

Relevant Courses: Algorithms, Machine Learning, Machine Learning in Finance, Database architecture and Design, Computation &Visualization, Data Mining in Engineering, Operations Research, Financial management

GVP College of Engineering, Visakhapatnam, India **Aug 2019**

Bachelor of Technology in Computer Science Engineering

Relevant Courses: Operating Systems, Statistics, Design and Analysis of Algorithms, Object Oriented Programming

TECHNICAL SKILLS

Programming Skills: Python, SQL programming, R, Java, JavaScript
Big Data: ETL, MATLAB, Hadoop, Spark, Hive

Libraries: NumPy, Seaborn, Pandas, Matplotlib, scikit-learn, TensorFlow

Tools: Tableau, Power Bi, Spacy, AWS, Git, NLT Toolkit, GCP, Microsoft azure, Microsoft excel, Redis.

Database: MYSQL, MongoDB,

Machine Learning: Supervised Learning, Unsupervised Learning, Deep Learning, Computer Vision, Natural Language Processing, Algorithms, Decision Trees, data Visualization, linear Regression, classification, Business intelligence, BI dashboard, data science, version control, problem solving, data structures, MLOPS.
Professional Certifications: TensorFlow developer by Google, Deep Learning specialization (Coursera)
Other Skills: Passionate, collaborative, attention to detail, excellent verbal and communication skills, team oriented, strong interpersonal skills

PROFESSIONAL EXPERIENCE

Data Scientist, Alma better. **May 2020-Apr 2021**

- Carried out performance evaluation employing Accuracy, RMSE, MAE and ROC to measure accuracy of models. And an accuracy of 85% is achieved on the validation.
- Created various ML models capable of detecting emotion through speech validated them, and back tested them.
- Implemented end to end using Flask API and deployed on AWS.
- Trained Deep Learning Models using Transfer learning in Python with help of TensorFlow.
- Performed analysis on audio clips using Mel Spectrogram, MFCC and Image analysis on the spectrogram of the audio clips using image processing techniques to make data- driven decisions.
- Remodeled using Data Augmentation methods and over sampling to create a robust model.
- Reduced hate-speech in virtual meetings by 87%.

PROJECTS

Stock Return Prediction, Northeastern University **Mar 2023**

- Conducted statistical analysis of a selected stock and its competitors by reporting summary statistics of the training period and plotting the kernel density, using Python.
- Constructed a feature database using various features/factors (FRED, Fama-French website, ADS, AR, CAPM, momentum factors, volume, price/return lags, etc.).
- Visualized the feature importance and feature selection process using regression-based approach (Ridge regression, LASSO, Elastic Net, or LARS) vs decision tree-based approach (random forest, Boost).
- Proposed and trained 3-6 models using the features prepared compared the model performance using RMSE between the fitted Y and actual Y in testing period, using Python.
- Composed trading rules that use buy-and-hold, long-short, or day trade.
- Generated trading signals using the above 3-6 models, compared their PNL, and implemented the best-performing model in a trading strategy.

Customer Revenue Prediction, Northeastern University **Feb 2022**

- Performed a Time Series and Geographic analysis on the data to gain insights and find patterns in the data.
- Used libraries like sklearn, seaborn in Python to Analyze a Google Merchandise Store (G Store) customer dataset as a part of Kaggle competition to predict revenue per customer.
- Trained models like Random Forest, XGBoost, KNN, Light GBM. After fine tuning hyperparameters, a consistent RMSE score of 1.71 has been achieved.
- Strengthened marketing ROI significantly by targeting audience using 80-20 rule and offered other financial services.

Sales Insight Dashboard, Tableau **Jan 2022**

- Created an interactive Tableau Dashboard to get insights into the revenue and Sales.
- Implemented a live connection to MySQL database.
- Generated charts like top n products, top n consumers filtered by year, Region.
- Discovered various patterns using data driven approaches in the shopping style of the customers that initiated a profitable target marketing.