MEGHAANA TUMMAPUDI

Boston, Massachusetts | (857) 376-9096
meghaanat.98@gmail.com | http://www.linkedin.com/in/meghaana | https://github.com/Meghaana

EDUCATION

Master of Science, Data Science — Expected: Dec 2023
Northeastern University, **Khoury College of Computer Sciences** — Boston, MA

Bachelor of Technology, Computer Science — June 2016 - Dec 2020
Jawaharlal Nehru Technological University — Hyderabad, India

TECHNICAL SKILLS

- **Programming Languages**: Python, SQL, R, C++.
- **Frameworks and Tools**: Keras, Tensorflow, PyTorch, PySpark, NLTK, Scikit-learn, OpenCV, Pandas, Numpy, MySQL, SQLite, PostgreSQL, dplyr, Tidyverse, AWS, Git, Plotly, Matplotlib, Seaborn, ggplot2, Tableau.
- **Data Science and Machine Learning**: Data Cleaning & Interpretation, Anomaly Detection, Dimensionality Reduction, Feature Engineering, Classification & Regression models.
- **Deep Learning**: CNN, RNN, LSTM, BERT based models, VGG, Densenet, Transformers.
- **Knowledge Domain:** Linear Algebra, Probability, Statistics, Data Structures, Algorithms, NLP, Information Retrieval, Image Processing, Supervised & Unsupervised ML.

WORK EXPERIENCE

Machine Learning Intern — Sept 2022 - Dec 2022
CodaMetrix — *Boston, MA*

- Analyzed over **1K+ medical reports** to understand prevalent issues and patterns, contributing to precise **report segmentation** and the effective resolution of problems.
- Improved **data quality** by **25%** via **pattern recognition** techniques, and **NLP** to correct missing and inaccurate data, resulting in a streamlined data pipeline.
- Devised and executed a training scheme that segmented medical reports into parts to generate additional data for an **SVM** model, resulting in a **12% accuracy boost** and a **33% automation rate** increase.
- Identified **potential areas** for **improvement** and handled **ad hoc requests** from various stakeholders to support their decision-making.

Graduate Teaching Assistant — Jan 2022 – May 2022
Northeastern University — *Boston, MA*

- Taught **Python**, **data analytics**, and **visualization** skills through DS2000 course, and provided comprehensive explanations of general **algorithmic techniques** in the CS3000 course.
- Managed class logistics and graded assignments for 200+ students, while collaborating with faculty and fellow TAs to provide one-on-one support, ensuring a seamless learning experience for students.

Data Science Intern — Jan 2021 - June 2021
Sabudh Foundation — *Hyderabad, India*

- Improved heart disease prediction **accuracy** by **20%** by utilizing **feature engineering** techniques, such as **Decision Trees (IG)** for **feature selection**, **data standardization**, and handling **class imbalance**.
- Created an SMS spam classifier with **97.3% accuracy** by implementing **Multinomial Naïve Bayes**, **Random Forests**, and word embedding techniques **(TF-IDF and word2vec)** while tackling **class imbalance** issues.

PROJECTS

Fake Check | **(Accuracy = 96.4%, F1-score = 0.95)**

- Built and trained a CNN-based image classification model, using **Vanilla CNN**, **VGG16**, **VGG19**, and **Densenet121** architectures, to differentiate real and fake human face images.

Question - Answering System | **(Accuracy = 70.9%, F1-score = 0.80)**

- Implemented Information Retrieval and BERT **(BERT, DistilBERT, ALBERT & Ensemble)** based models to create a Question-Answering Model on Stanford Question & Answering (SQuAD 1.1) dataset.

Sentiment Analysis | **(Accuracy = 72.15%)**

- Developed a sarcasm detection system on Reddit comments using RNNs like **Bidirectional LSTM** in combination with word embedding techniques **(TF-IDF, GloVe, and fastText).**

Binary Classification on UW Breast Cancer Data | **(Accuracy = 93%)**

- Built binary classifiers utilizing **Linear Discriminant Analysis**, **Quadratic Discriminant Analysis**, and **Logistic Regression** on the UWBreast Cancer dataset to classify cysts as malignant or benign.