# Naroju Vamshidhar

+919502558273 | narojuvamshidhar@gmail.com | linkedin | github

## Education

**University of Hyderabad** — Hyderabad, India
*Master of Technology in Computer Science with 8.03 CGPA* — *Aug. 2018 – June 2020*

**Matrusri Engineering College** — Hyderabad, Telangana, India
*Bachelor of Engineering in Computer Science and Engineering with 81 percentage* — *Aug. 2014 – May 2018*

## Experience

**Teaching Assistant** — June. 2019 – June 2020
*University of Hyderabad* — *Hyderabad, India*

Responsibilities

- Teaching students of Integrated MTech and MCA Virtualization and Machine Learning concepts
- Implementing Machine Learning and Deep Learning algorithms on various real-world datasets

## Projects

**Predicting Project Approval in DonorChoose.org** | *Python, K-NN, Naive Bayes, DecisionTree, GBDT, LSTM*

- DonorsChoose.org receives hundreds of thousands of project proposals every year for classroom projects in need of funding. Right now, a large number of volunteers are needed to manually screen each submission before it's approved to be posted on the DonorsChoose.org website.
- The project is about automating this manual screening process and predict the project approval by applying machine learning algorithms and performing some Natural Language Processing Tasks
- github code LSTM ML

**Amazon Fine Food Review Analysis using NLP Algorithm BERT** | *Keras, TensorFlow, BERT*

- Sentiment Analysis of the review given by the user by performing Natural Language Processing Tasks
- Used Bidirectional Encoder Representation from Transformers with some Neural Network Layers on top of it
- Performed Data Preprocessing, Tokenization of sentences and used BERT Embeddings
- Successfully computed the sentiment of the review using the BERT and Deep Learning Architecture with good accuracy
- github code link

**Human Activity Recognition Using Smart Phones** | *Logistic Regression, LinearSVC, Random Forest, LSTM*

- Dataset for this problem is obtained from the University of California Irvine (UCI) Machine Learning Repository
- The database is built from the recordings of 30 subjects performing activities of daily living while carrying a waist-mounted smartphone with embedded inertial sensors
- Applied Machine Learning Algorithms like Logistic Regression, LinearSVC and RandomForest with experts extracted features
- Applied Deep Learning Algorithm LSTM on raw time series data
- github code link

**Classification of Documents using CNN** | *TensorFlow, Keras, CNN, Word-Embedding, Character-Embedding*

- This project deals with documents of 20 different types
- Mapping the documents to the respected class label is the major challenge to be solved
- Performed Exploratory Data Analysis, Data Cleaning, and Data Preprocessing
- Used word embedding with the pre-trained glove models for one of the models and character embedding for the other
- Built a complex Neural Neural architecture to solve the task
- github code link

**Social Network Link Prediction** | *Python, Scikit-learn,NetworkX,EDA*

- The task in this project is about predicting the missing links in the directed social graph and recommending possible connections to users

- The dataset for this project is obtained from Kaggle facebook's recruiting challenge
- Performed Exploratory Data Analysis, Feature Extraction, and Hyper-parameter Tuning on the given dataset
- Applied Machine Learning Algorithm XGBoost for predicting the link between social network users
- github code link

**Microsoft Malware Detection** | *KNN, Scikit-learn,Logistic Regression, XGBoost, Random Forest,t-SNE*

- The major part of protecting a computer system from a malware attack is to identify whether a given piece of file or software is a malware
- There are total 10,868 .bytes files and 10,868 .asm files total 21,736 files in the given dataset. These files play an important role in detecting malware
- Performed Feature Extraction and Hyper-parameter Tuning
- Applied different Machine Learning Algorithms for predicting the malware-class
- github code link

**DenseNet on CIFAR-10** | *Python, CNN, Keras, TensorFlow, Scikit-learn*

- The Convolutional Neural Network used in this project is from the research paper Densely Connected Convolutional Networks
- Based on research paper link
- Trained the network on the CIFAR-10 dataset using Image augmentation techniques
- github code link

**Question pair similarity matching in Quora** | *Python, Scikit-learn, EDA, LogisticRegression, LinearSVM, XGBoost*

- Quora is a place to gain and share knowledge about anything. It's a platform to ask questions and connect with people who contribute unique insights and quality answers
- The task in this project is to match the questions that have the same intent which helps both the readers and the writers
- Performed Exploratory Data Analysis, Data Preprocessing, and Data Cleaning on the given dataset
- Performed Feature Extraction and Hyper-parameter Tuning on the given dataset
- Applied different Machine Learning Algorithm like Logistic Regression, Linear SVM, XGBoost for finding the similarity of questions
- github code link

## Technical Skills

**Languages**: Python, C/C++, MySQL
**Libraries and Framework**: Scikit-learn, Pandas, Numpy, Scipy, Keras, Matplotlib, NLTK, TensorFlow, Jupyter Notebook, Flask
**Areas of Interest**: Machine Learning, Deep Learning, Natural Language Processing, Neural Networks, Transformer models, Generative Adversarial Networks, Computer Vision, Product Development