Saideep Reddy Pakkeer

9256 E, Regents Road, La Jolla, California 92037

646-250 2544 | saideepreddy91@gmail.com | saideepreddy91 | saideep-reddy

Education

University of California San Diego *Anticipated Mar. 2020*

- Masters in Machine Learning and Data Science GPA - 3.96/4.0

Indian Institute of Technology, Bombay (India) *2012 - 2016*

- Bachelor of Technology with honors in Electrical Engineering GPA - 8.23/10

Experience

Data Scientist Intern - Intel Corporation, Phoenix, Arizona Area *Jun. 2019 - Present*

Data Analytics and Statistical Methods Team

- Working on optimizing the process control systems across teams (module engineers) using NLP
- Working on building unsupervised models to cluster 3-D shapes using autoencoders

Analyst, Data Science - Actify Data Labs, Bangalore, India *Nov. 2017 - Aug. 2018*

Sentiment Analysis

- Built an end-to-end pipeline in **Django** (dashboard, playback, upload among other functionalities) to score an audio signal for sentiment using an ensemble of **gradient boosting** models
- Classified speakers according to his/her identity using **Hidden Markov Model** on the audio signal (Speaker Diarization)
- Designed the web framework(**MVC architecture**) and also achieved an accuracy of **85%** on the ML model for binary sentiment

Cancer nodule detecting system

- Developed cancer nodule detection system using **mask R-CNN** implementation on DICOM images using **Tensorflow**
- Trained & tested on the Digital Database for Screening Mammography(DDSM) and worked with a local hospital in India for beta testing

Analytics Specialist - Opera Solutions, Noida, India *Jun. 2016 - Oct. 2017*

Identifying tax evaders, Operation Clean Money - Govt. of India

- Designed a likelihood model (**Logistic Regression & XGBoost**) for the high-profile demonetization project for identifying tax-evaders likely to respond to Govt. notices working directly out of the **Income Tax Department**
- Engineered a predictive feature set from **huge** & **diverse** data sources - Income tax **returns**, bank **transactions**, property purchases

Predicting cargo booking weight

- Built ensemble of gradient boosting models (**xgboost**) for predicting cargo show-up rate for a major airline
- Tuned the ensemble model (eta, max depth, min child weight) for accurate prediction of the shipments tendered weight within 5 percent error range for **96%** of the bookings and **deployed** the model

Research and Projects

Multi-label classification of news articles | NLP & Recommender Systems *Sep. 2018 - Jun. 2019*

- Implemented a text classifier for categorizing news articles into 30 categories (crime, health, sports) using a **DenseNet** neural network
- Demonstrated the performance among different classifiers using techniques like **TF-IDF, n-gram** and achieved an accuracy of **96.8%**
- **Deployed** the model and built a web application in **Dash** {https://cse256.herokuapp.com/}

Multi-class image classification on Fashion MNIST | Deep Learning *Sep. 2018 - Dec. 2018*

- Built different classifiers (**ResNet, VGG, LeNet**) using different architectures of **CNNs** to classify Fashion MNIST images (10 classes)
- Using ensembling techniques to boost weak learners and make a strong and robust model achieved an accuracy of **94.3%**

Linear program for non-convex function approximation | Convex Optimization *Dec. 2015 - May. 2016*

- Developed a linear program for approximating a non-convex function with a **convex envelope** which is a major concern for many machine learning problems in recent years (**Bachelor Thesis**)
- Demonstrated the performance on various non-convex functions, incorporting the ideas of linear function approximation and **constraint sampling** to reduce the curse of dimensionality by simulating the linear program in MATLAB

Publication

Approximating convex envelopes using linear programming *Nov. 2018*

- Developed a linear program using Oberman's characterization of convex envelope for approximating any non-convex function with a convex envelope and submitted to the journal **Annals of Operations Research** (ANOR-D-16-01198)

Achievements & Awards

2012 **All India 91st rank** in Indian Institute of Technology - Joint Entrance Exam among 500,000 students *National*

2012 **All India 26th rank** in AIEEE (All India Engineering Entrance Examination) among 1 million students *National*

Skills & Courses

Languages & skills: • Python • SQL • R • Matlab • Spark • Django • Dash • Git • Sklearn • Tensorflow/PyTorch • AWS

Relevant courses: • Statistical Natural Language Processing • Deep Learning for Computer Vision (CNNs, Auto-Encoders, Few-shot Learning, GANs) • Recommender Systems & Web Mining • AI: Learning Algorithms • Data Analysis using Spark

Positions of Responsibility

- **Graduate Teaching Assistant** in the Math department at UCSD - Winter'19 & Spring'19 • **Computer Science Tutor** at UCSD - Fall'18
- **Alumni Secretary**, Electrical Engineering Dept.: Conducted Student Alumni Meet with the Alumni Relation Cell at IIT Bombay