

Saideep Reddy Pakkeer

☎ 646-250 2544 | ✉ saideepreddy91@gmail.com | 📱 saideepreddy91 | 🌐 saideep-reddy

Education

University of California San Diego

- Masters in Machine Learning and Data Science

May 2020

GPA - 3.97/4.0

Indian Institute of Technology, Bombay

- Bachelor of Technology with honors in Electrical Engineering

2012 - 2016

GPA - 8.23/10

Experience

Data Scientist Intern - Intel , Phoenix, Arizona

Jun. 2019 - Sep. 2019

PATTERN RECOGNITION

- Built **predictive models** using NLP techniques **working across** the module engineering **teams** to identify important factors affecting test failures for different products
- Identified the primary reasons for class test failures and recommended the remedy steps to decrease downtime of the tools
- **Automated the python scripts** to run on newer data in the future for production and later to be integrated to internal systems at Intel
- Built unsupervised deep learning models (autoencoders) to **cluster 3-D shapes** from the coplanarity measurements on the substrate
- Learnt low dimensional representation of high dimensional coplanarity data to identify defects due to solder bridging

Data Science Analyst - Actify Data Labs (AI Startup), Bangalore, India

Nov. 2017 - Aug. 2018

SENTIMENT ANALYSIS

- Built an end-to-end pipeline in **Django** (dashboard, playback, upload among other functionalities) and pushed it to production to classify an audio segment for sentiment using an ensemble of gradient boosting models with more than 85% accuracy
- Classified speakers according to his/her identity using **Hidden Markov Model** on the audio signal (**Speaker Diarization**)

CANCER NODULE DETECTING SYSTEM (Object Detection)

- Developed cancer nodule detection system using **mask R-CNN** implementation (transfer learning) on DICOM images in **Tensorflow**
- Trained & tested on the Database of Mammography (DDSM) using **Deep Learning API** working with a local hospital 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) to send **targeted emails/notices** 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
- **Collaborated** with income tax officials on a daily basis to design the models and understand the datasets

PREDICTING CARGO BOOKING WEIGHT

- Built ensemble of gradient boosting models (**xgboost**) for predicting cargo show-up rate (overbooking vs no-show) for a major airline
- Achieved accurate prediction of the shipments tendered weight within 5 percent error range for **96%** of the bookings and **deployed** the model built entirely in python

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
- Ranked 9/171 (**top 6%**) in the **Kaggle** competition on sentiment analysis (NLP class)
- Demonstrated the performance among different classifiers using techniques like **TF-IDF, n-gram** and achieved an accuracy of **96.8%**
- Built a web application and **deployed** the model 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 (**Bachelor Thesis**)
- Demonstrated the performance on various non-convex functions, incorporating the ideas of linear function approximation and **constraint sampling** to tackle the curse of dimensionality by simulating the linear program in MATLAB

Publication

Approximating convex envelopes using linear programming

Nov. 2018

- *Akhil Shetty, Saideep Reddy, Vivek S. Borkar, Neeraja Sahasrabudhe*, submitted to **Annals of Operations Research (ANOR)**

Skills & Courses

Languages & skills: • Python • SQL • Spark • Scikit, Pandas • Pytorch • Tensorflow/Keras • Django, Dash • AWS • Git • Matlab • R • Data Visualization - Matplotlib, Bokeh, Altair, Ipywidgets • JMP (Statistical software)

Relevant courses: • Statistical Natural Language Processing (Language Modelling, Machine Translation, Sequence Tagging, Text Classification) • Deep Learning for Computer Vision • Design of experiments^{Intel} • Recommender Systems & Web Mining (Latent-factor models, Collaborative filtering) • AI: Learning Algorithms • Big Data Analytics Using Spark

Achievements & Positions of Responsibility

- All India **91st rank** in Indian Institute of Technology - Joint Entrance Exam among 500,000 students - 2012
- All India **26th rank** in AIEEE (All India Engineering Entrance Examination) among 1 million students - 2012
- **Graduate Teaching Assistant**, for 4 quarters in the Math department at UCSD
- **Computer Science Tutor** at UCSD - Fall'18
- **Alumni Secretary**, Electrical Engineering Dept.: Conducted Student Alumni Meet with the Alumni Relation Cell at IIT Bombay