

Saideep Reddy Pakkeer

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

Education

University of California San Diego

- Masters in Machine Learning and Data Science

June 2020

GPA - 3.98/4.0

Indian Institute of Technology, Bombay

- Bachelor of Technology with honors in Electrical Engineering

2012 - 2016

GPA - 8.23/10

Experience

CV Performance Analysis Engineer - Apple

July, 2020 - Present

COMPUTER VISION MACHINE LEARNING

Bay Area, California

- Working in the CVML team

Data Scientist Intern - Intel

Jun. 2019 - Sep. 2019

PATTERN RECOGNITION & SHAPE ANALYSIS

Phoenix, Arizona

- Built **predictive models** using Natural Language Processing (NLP) based techniques **working across** the module engineering **teams** to identify important factors affecting test failures for different products
- **Automated python scripts** to generate reports & built a pipeline to run on future data

Data Science Analyst - Actify Data Labs (AI Startup)

Nov. 2017 - Aug. 2018

SENTIMENT ANALYSIS

Bangalore, India

- 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 & Localization)

- Developed cancer nodule detection system using **mask R-CNN** implementation (transfer learning) on DICOM images in **Tensorflow**
- Trained & tested on the combination of public and clinic specific data using **Deep Learning AMI** working with a hospital

Analytics Specialist - Opera Solutions

Jun. 2016 - Oct. 2017

IDENTIFYING TAX EVADERS, OPERATION CLEAN MONEY - GOVT. OF INDIA

Noida, India

- Designed a likelihood model (Logistic Regression & XGBoost) to send **targeted emails/notices** for identifying tax-evaders
- 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 working out of the Income tax dept.

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

Research and Projects

Real-time multi-label classification | 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%**

Gun Violence in the US | Web Scraping and Interactive Visualization

Jan. 2019 - Mar. 2019

- Scraped the Gun Violence online archive using **Beautiful Soup** package to create the dataset
- Built **interactive tools** using Bokeh, Seaborn & other packages and did an exploratory data analysis on gun violence across different states to identify the relation between gun laws, mental health and suicide rate across different states

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) • A/B testing

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) • Statistical Learning (Bayesian, ML, EM) • Big Data Analytics Using Spark (Docker, 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** (150 students), for 4 quarters - Math department, 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