# Saideep Reddy Pakkeer

| saideepreddy91@gmail.com | saideepreddy91.github.io | saideepreddy91 | saideep-reddy

## Education

**University of California San Diego** *June 2020*

- Masters in Machine Learning and Data Science GPA - 3.98/4.0

**Indian Institute of Technology, Bombay** *2012 - 2016*

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

## Experience

**Machine Learning Engineer - Apple** *Aug. 2020 - Present*

COMPUTER VISION MACHINE LEARNING *Sunnyvale, California*

- Working on data related aspects to support algorithm development and evaluation using python, Xcode, Matlab, and computer vision tools like OpenCV
- Working on Data manipulation and visualization tools and libraries in python like pandas, matplotlib
- Balancing research and product to deliver Apple quality, state-of-the-art experiences, and innovation through the full stack
- Partnering with Hardware, Software, and Machine Learning teams to influence the sensor and silicon roadmap that brings the Apple vision to life
- Evaluate neural network algorithms in TensorFlow and Keras
- Working on bringing up the data infrastructure tools for storing and managing data in the Video Engineering 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

**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*

PREDICTING CARGO BOOKING WEIGHT *Noida, India*

- 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%**

## 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