7675 Palmilla Dr, Apt 6106
San Diego, CA-92122

MITESH GADGIL

Github: miteshgadgil
Linked In: miteshgadgil

Phone: +1 (858)-729-8137 **Webpage**: https://miteshgadgil.github.io **Email**: mgadgil@ucsd.edu

Objective

Seeking a **full-time** position where software skills can be utilized for **Data Analysis & Machine Learning**

Education

UC San Diego **San Diego, CA** **Expected March '17**
Masters in ECE (Major: Intelligent Systems) **GPA**: 3.58/4.0

BITS Pilani **Goa, India** **June '15**
Bachelor of Engineering (Hons.) in Electrical & Electronics **GPA**: 8.82/10.0

Skills

Programming and Tools: Python, R, Apache Spark, SQL, MATLAB, C, Tableau
Modelling Techniques: Regression, Classification, Clustering, Ensembling, Boosting, Stacking and Blending

Relevant Coursework

- Exploratory Data Analysis & Inference using R
- Big Data Analytics using Spark
- Statistical Learning
- Information Visualization* (*currently enrolled)

Work Experience

Graduate Teaching Assistant **University of California, San Diego** **June '16 - Present**
- Conduct weekly physics labs involving **data collection** and **error analysis** in MATLAB for a class of 30 students
- Communicate complex signal processing concepts to a class of 100 students, in an easy to follow manner

Software Engineering Intern **Tonbo Imaging Pvt. Ltd., India** **Jan – June '15**
- Designed experiments and a data collection method to evaluate the performance of 2 temperature sensors
- Analysed the experimental data using Excel and drafted a **data-driven recommendation** for the VP Engineering

Software Engineering Intern **Mapyn Technologies Pvt. Ltd., India** **Mar – May '13**
- Conceptualized 'smart' safety features for a motorized platform used to lift industrial loads
- Implemented a **linear regression** model which predicted weight of the load using current drawn by the motor

Projects

Loan Granting Binary Classification **August '16**
- Built a model to predict whether an applicant will repay or default on a loan based on data about the applicant
- Formulated the **end-to-end modelling** process for the data(255,000 rows & 19 features) using **IPython notebook**
- Achieved an accuracy of **84.7%** using Random Forest model and investigated feature importance from the model

Analysing Twitter Data using Spark **May '16**
- Analysed the 10 most frequently used tokens in tweets of various user groups to infer their opinions and beliefs
- Wrote a script using **PySpark** that deployed an Amazon cluster to mine **20 GB** of twitter data in under 2 minutes

Predictive modelling for Insurance Claim Approvals (Kaggle Contest) **March '16**
- Trained a model which predicts the probability that an insurance claim can be expedited for approval
- Performed feature exploration and selection on the anonymized data; trained an **XGBoost** model in **R**
- Achieved 1298th (**Top 45%**) place among 2900 participants in my first ever Kaggle contest