# PARIN SHAH

Los Angeles, CA|psshah@usc.edu|(213) 285-8984|github.com/Parin09|linkedin.com/in/parin-s-shah|parin09.github.io/portfolio/|Tableau

## EDUCATION

**University of Southern California, Los Angeles**           **Aug 2021-May 2023**
Candidate for Master's in Computer Science (Data Science)
Relevant Coursework: Analysis of Algorithm, Foundations of Artificial Intelligence

**Mukesh Patel School of Technology Management and Engineering, NMIMS, Mumbai**     **Aug 2016-May 2020**
Graduated as Bachelor of Computer Engineering with Distinction (Merit List)         **83.5%**
Relevant Coursework: Data Structures, Database Management Systems, Data Warehousing and Mining, Artificial Intelligence, Business Viz

## PROFESSIONAL EXPERIENCE

**Limechat, Mumbai | Data Scientist**           **Jun 2020-Jun 2021**

- Devised a tool to label new intents using **clustering** and **active learning** reducing manual labeling time by over **60%**.
- Managed **6** cross functional projects by providing project leadership and daily management throughout the project. Some of these projects are Customer Satisfaction Score, Data Labelling, Re-Engagement campaigns.
- Examined drop-off of users at various stages to target certain groups in **Re-Engagement campaigns** leading to an increase in sales by **28%** on average across companies.
- Engineered a **heuristic-based** algorithm to calculate **Customer Satisfaction Score** for each chat. Handed off unsatisfactory chats to customer support improving the overall customer experience.
- Designed and implemented a real-time data pipeline in python to process and upload over **50,000+** rows of semi-structured data per day from PostgreSQL server to product analytics tool (**Mixpanel**).
- Communicated weekly insights to over **70+** stakeholders to help take an informed decision on changes and future features.
- Developed a custom analytics dashboard on Flask and VueJS displaying critical **KPIs** for each client with ability to download standard excel and pdf reports. Integrated caching algorithm using Redis as cache that led to **90%** faster loading times.
- Technologies used: Python, SQL, Mixpanel, Flask, VueJS, Azure, Gitlab, Postgres, Redis

**Birthvenue, Mumbai | Project Intern**           **Mar 2019-May 2020**

- Developed a universal rating platform for all types of cryptocurrencies and tokens available in the market based on financial and non-financial variables to help investors make decisions.
- Designed a **regularized regression** model that takes in variables, determines the rating and displays them for **1100+** cryptocurrencies on website.
- Technologies used: Python, Flask, ReactJS, MySQL, AWS.

**Oracle Financial Services Software Limited, Mumbai | Research Intern**           **May 2019-Jul 2019**

- Interlinked **LDAP** server in Kubernetes for authorization purposes. Automated manual formation of access roles in banking systems helping **save 10+** hours of work per client.
- Incorporated **OpenID** connect to fetch relevant roles from server. Authorized multiple users by assigning permission to roles with the help of **Role Based Access Control (RBAC)**.

## ACADEMIC PROJECTS AND PAPER

**Immersive Visualization in Medical Imaging:** Reports 3D (Paper) | Python, Blender, Flask, VTK, Keras

- Constructed a web application to bridge the semantic gap between medical practitioners and laymen by leveraging the use of **Augmented Reality** (AR) as a graphically intensive solution.
- Devised novel algorithm for perfect 3D **volumetric recreation** and **rendering** of organ. Generated 3D file of the organ with tumor to be viewed in AR.
- Achieved a **DICE** score of **87%** for segmenting HGGs (Brain Tumours) from MRI by applying U-Net CNN architecture.

**Review of Credit Card Fraud Detection Techniques** (Paper) | Python, Numpy, Sci-Kit Learn

- Implemented **Synthetic Minority Oversampling Technique** to generate synthetic dataset to improve size of minority class.
- Employed **logistic regression** to classify transactions and achieved a F1 score of **94%**, Recall of **99%** and Precision of **90%**.
- Analyzed currently existing credit card fraud detection algorithms based on various parameters such as pre-processing, complexity, computation time, accuracy and listed the advantages and disadvantages with a suitable use-case.

**Grocery Store Case Study** | Alteryx, Excel, Tableau

- Provided recommendation by applying various analytical techniques on how to expand a grocery store chain.
- Clustered existing stores using **K-Means**. Predicted clusters for new stores applying **Boosted Model** with accuracy of **83%**.
- Forecasted fresh produce sales of every month for the next year based on historical data by employing **ARIMA** model.

## TECHNICAL SKILLS

- Programming Languages and Databases: Python, SQL, SAS, C++, Postgres, Redis, MySQL
- Python Libraries: Matplotlib, Plotly, Pandas, Numpy, Scipy, Tensorflow, Keras, Sci-Kit Learn
- Software and Frameworks: Azure, AWS, GitLab, Tableau, Alteryx, Docker, Kubernetes, Jira, Notion
- Web Technologies: Flask, Django, VueJS, ReactJS

## LEADERSHIP AND INVOLVEMENT

- **Technical Head, Association for Computing Machinery (ACM) MPSTME -** conducted and taught in technical workshops including AR on Unity 3D, C programming, and served as a **project mentor** for teams working in the domain of AR.
- **Android Developer**, Rotaract Club of Bombay Uptown - created informatory PR application having **4.9** stars on play store.