

## EXPERIENCE

**Software Engineer (Data Team)** [MerQube, Inc](#) San Francisco, USA July 2023 — Present

- Engineered the migration of equity reference and end-of-day pricing pipelines to a new provider platform, ensuring uninterrupted data delivery for index calculations with zero downtime.
- Led the development of a scalable options data platform, defining the data model, building ingestion and monitoring systems, and partnering with product, data providers, and financial engineers to resolve complex data integrity challenges. Delivered a unified data access layer that powers multi-asset index development and self-service analytics across teams.
- Transformed operations workflow by developing automated validation checks for end-of-day prices and corporate action data across multiple providers. Enhanced data accuracy and consistency with customizable override capabilities.

**Perception Software Intern (Acoustic ML Team)** [Zipline International](#) South San Francisco, USA June 2022 — Aug 2022

- Performed in-depth analysis of a vast flight data corpus to uncover spatial and temporal correlations, successfully identifying likely sources of false positives between airborne and ground-based objects.
- Fine tuned the intruder detection models by strategically integrating hard negative samples, leading to a decrease in false positives while preserving optimal sensitivity.

**Data Scientist** [Karza Technologies Pvt. Ltd.](#) Mumbai, India Sept 2018 — Mar 2020

- Successfully delivered an end to end OCR Pipeline for the KYC Documents by implementing a robust synthetic data generation pipeline, model validation, hypothesis testing and training modules for text recognition, card detection and text detection tasks.
- Developed a light weight reformulation of fixed length text recognition model using the CTC framework by progressively optimizing the model towards bridging the latency vs task specific evaluation metric trade-off.
- Engineered the first deployment of deep learning models across AWS Lambda serverless functions and GPU instances, meticulously optimizing for compute, memory, and response time requirements.

## PROJECTS

**Cartoonizer** (Featured in [TheNextWeb](#) and multiple other blogs) May 2020 — Jan 2021

- Developed a Generative Adversarial Network model to convert images into a cartoonized style and deployed it as an end to end serverless web application [Cartoonizer](#) on Google Cloud Run.
- Optimized the inference pipeline for videos and deployed the model by packaging it as a containerized application using docker and hosting that container in a serverless environment on Google Cloud Run and Algorithmia.
- The application had been used by over 200k users and even cartoonized a short movie ([The Black Disquisition \(Video\)](#)), selected at multiple film festivals and won award for its animation too.

**Understanding of Effective and Emotional Components in Advertisements**

Sept 2017 — Apr 2018

- Defined the emotional components in the form of arousal and valence scores using the tagged sentiment of each ad and mapping each sentiment to its corresponding score using a score dictionary obtained by training on a large text corpus.
- Created an effective score as a proxy metric using custom heuristics based on the user responses to the questions "What does the ad tell you to do?" and "Why does the ad tell you to do it?".
- Trained and fine tuned the DenseNet-121 neural network, with individual models to predict arousal, valence and effectiveness scores for [image advertisements](#).
- Paper was accepted as a poster presentation for the CVPR Understanding Visual Advertisements workshop, Salt Lake City, USA, 2018.

## EDUCATION

**Master of Science in Computer Science**, New York University (Courant)

May 2023

GPA: 3.8/4.0

Relevant Courses: Deep Learning (TA for 2 semesters), NLP, Computer Vision, High Performance Machine Learning.

**Bachelor of Engineering in Computer Engineering**, University of Pune (Maharashtra Institute of Technology)

May 2018

GPA: 3.8/4.0

Relevant Courses: Operating Systems, Algorithms, Distributed Computing, Database Management, Data Mining.

## TECHNICAL SKILLS

**Programming Technologies**

Python, C++, Java, MySQL, Postgres SQL, MongoDB (NoSQL)  
Tensorflow, Pytorch, Numpy, Pandas, Spark, Airflow, AWS (Sagemaker, Redis, Lambda Functions, ECS, Batch) GCP (Cloud Functions, Cloud Run), Git, Jupyter, Docker, Metaflow, Kubernetes, TensorRT, Ray.