

# Malay Koladiya

510-320-8438 | [malaykoladiya.01@gmail.com](mailto:malaykoladiya.01@gmail.com) | [linkedin.com/in/malaykoladiya/](https://www.linkedin.com/in/malaykoladiya/) | [github.com/malaykoladiya](https://github.com/malaykoladiya) | Portfolio

## Education

### California State University, Fullerton (GPA - 3.77/4.00)

Fullerton, CA

Master of Science in Computer Science

Jan 2022 – May 2024

- **Relevant Coursework:** Advanced Algorithms, Web Backend Engineering, Advanced Database Management Systems, Introduction to Machine Learning, Advanced Blockchain Technologies, Artificial Intelligence, Data Structures and Algorithms, Advanced C++

### San Jose State University (GPA - 3.87/4.00)

San Jose, CA

Bachelor of Science in Aerospace Engineering, Minor in Computer Science

Jan 2017 – May 2021

## Technical Skills

**Languages:** Python, JavaScript, C++, Rust, C, Core Java

**Databases:** MongoDB, Redis, PostgreSQL, MySQL

**Frameworks/Libraries:** Django, Flask, Next.js, Express.js, React.js, Node.js, FastAPI, Socket.io, Docker, Kubernetes, HTML5, CSS3, Tailwind CSS, Actix-Web, Git, pandas, NumPy, Matplotlib

**Cloud Services:** Amazon Web Services (EC2, S3, Fargate, Lambda, API Gateway)

**Big Data Tools:** Hadoop, Apache Spark, MapReduce, Apache Kafka, Apache ZooKeeper

## Experience

### Graduate Research Assistant

Jan 2023 – May 2023

California State University, Fullerton (in partnership with Digiclips Inc)

Fullerton, CA

- Led the integration of a Video Text Recognition Docker container, optimizing Free-To-Air Antenna Television data processing pipeline
- Engineered an automated ETL (Extract, Transform, Load) workflow reducing video file ingestion and processing time by 25%, and ensuring seamless integration within a containerized environment.
- Designed, configured, and maintained a MySQL database schema to store and query over 100,000 textual data records extracted from video sources
- Diagnosed and resolved critical issues within the speech-to-text container, achieving a 15% reduction in downtime and ensuring uninterrupted text extraction
- Created detailed technical documentation of Docker container API endpoints and internal architecture to support continuous development and cross-team collaboration

## Projects

### JobsNearby – Job Search Portal | *Flask, React, AWS Lambda, API Gateway, Docker, Redis*

Dec 2023 – Apr 2024

- Architected a job search platform with user authentication, job listings, and profile management features for employers and applicants
- Designed a highly scalable system utilizing AWS Lambda and API Gateway, with MongoDB and Redis integration for efficient data management and high-performance operations
- Enhanced platform performance by implementing advanced lazy loading techniques, effective caching strategies, and optimizing unused CSS/JS, resulting in a Lighthouse performance score of 81

### Real-Time Chat Application | *MERN, Socket.io, Daisy UI, Tailwind CSS*

Jan 2024 – Feb 2024

- Developed a real-time communication platform with live messaging and user status using Socket.io within a MERN architecture, and a React and Tailwind CSS front-end for seamless user experiences
- Facilitated JWT-based authentication and designed MongoDB schemas to safeguard and manage user data effectively
- Implemented real-time notifications and optimized WebSocket connections for high concurrency and minimal latency

### Real-time Stock Market Data Streaming | *AWS EC2, Databricks, Kafka, ZooKeeper*

Jan 2023 – May 2023

- Designed a real-time stock market data streaming pipeline using Apache Kafka clusters hosted on AWS EC2 instance, integrating data from Yahoo Finance and Kaggle for user-driven investment decisions
- Utilized PySpark and Spark Structured Streaming for real-time data processing and analysis within the Databricks environment, building interactive dashboards to visualize market trends and metrics

### Wordle Web Backend Application | *Quart, Nginx, Redis, SQLite*

Aug 2022 – Dec 2022

- Developed Wordle's event-driven backend infrastructure and RESTful APIs, incorporating a real-time leaderboard and ensuring seamless user experience
- Implemented a scalable and fault-tolerant database system using SQLite with data replication, ensuring high data integrity and availability
- Optimized system resilience and performance with Nginx for HTTP reverse proxy and load balancing, and Redis Queue for enhanced message queuing between decoupled applications and webhooks integration for microservices event handling

### Myers-Briggs Personality Test Indicator | *Pandas, NumPy, NLTK*

Jan 2022 – May 2022

- Implemented and evaluated performance of 4 machine learning classifiers (SVM, Random Forest, K-NN, and Logistic Regression) to accurately predict MBTI personality types from textual data
- Conducted data pre-processing and feature engineering on dataset of 10,000 rows, incorporating techniques such as Lemmatization, Vectorization, and TF-IDF transformation to enhance model input quality
- Leveraged a comprehensive stack of data science libraries, including Pandas, NumPy, Matplotlib, NLTK, Scikit-Learn, and Imbalanced-Learn, to construct, train, and validate models, employing metrics such as confusion matrices, precision, recall, and F1-score to determine classifier effectiveness