

KEVIN CHUANG

SOFTWARE ENGINEER · DATA ENGINEER

🏠 San Jose, CA
📧 k-chuang.github.io

☎ (408) 691-1788
📧 github.com/k-chuang

✉ kevinchuang7@gmail.com
🌐 linkedin.com/in/kevin-chuang

WORK EXPERIENCE

Turo

DATA ENGINEER

San Francisco, CA

09/2023 - Present

- Built data pipelines with data quality checks for multi-channel, multi-touch marketing attribution, enabling marketing and analytics teams to better allocate spend and explore growth opportunities using **Python, Airflow, SQL, Segment, AWS (Redshift)**
- Contributed to core data services framework by creating reusable, tested software components for ETL, data quality, and data profiling using **Python, SQL**

HomeLight

SENIOR DATA ENGINEER

San Francisco, CA

07/2022 - 04/2023

DATA ENGINEER

11/2020 - 07/2022

- Designed, built, and launched core algorithm to match and rank real estate agents with home buyer and seller opportunities, **boosting conversion rate by 15%** using **Python, SQL, Django, ElasticSearch, Airflow, AWS (Kinesis, S3, Redshift), Heroku**
- Developed internal TV ad planning and bid optimization system to support marketing team, **resulting in reduction of \$250,000 annually in external ad agency costs** using **Python, Django, Airflow, SQL, AWS (ECS, S3, RedShift), Heroku**
- Spearheaded design and deployment of mission-critical data pipelines to showcase property listing images and transaction history, **enhancing user experience on company's primary website** using **Python, Airflow, SQL, ElasticSearch, Kafka, AWS (Lambda, SQS, S3, Redshift, Kinesis, DynamoDB)**
- Implemented record linkage solution to support operations team in tracking real estate agents and transactions, **leading to discovery of missed revenue opportunities totaling over \$500,000** utilizing **Python, SQL, Airflow, Redshift, ElasticSearch**

ViacomCBS

MACHINE LEARNING ENGINEER

San Francisco, CA

02/2020 - 11/2020

DATA SCIENCE INTERN

06/2019 - 12/2019

- Designed, built, and launched machine learning application to automatically generate and expose engaging video previews to support internal brands (**CBS News, CBS All Access**) and remove dependency on external services, **resulting in significant cost savings** using **Python, Django, Kubernetes / Helm, GCP (GKE, Pub/Sub, GCS, Cloud SQL)**
- Built serverless collaborative filtering recommendation system to power personalized movie / show recommendations for **CBS All Access** using **Python, Spark, Redis, Kubernetes, Argo** and **GCP (BigQuery, Dataproc, GKE)**
- Developed and launched backend API for creating personalized movie / show recommendation carousels for millions of new & trial users on **CBS All Access**, **enhancing user experience** using **Python, Django, Redis, Kubernetes, GCP (GKE, Cloud SQL)**
- Researched various text summarization methods (TextRank, Pointer Generator) to create cloud-native text summarizer application consisting of Chrome extension (**JavaScript**) with **Python Flask** backend server deployed on **Google App Engine** using **Docker**

Knowles Intelligent Audio

DATA ENGINEER

Mountain View, CA

02/2017 - 05/2019

TEST ENGINEERING INTERN

10/2016 - 02/2017

SKILLS

Programming

Python, SQL, Bash, Golang, JavaScript, Java, MATLAB

Database

PostgreSQL, RedShift, DynamoDB, BigQuery, MySQL, SQLite, MongoDB, Redis

Tools

Airflow, Django, ELK (Elasticsearch, Logstash, Kibana), Pandas, Spark, Git, Docker, Kubernetes, Helm, Pytest, NumPy, SciPy, Matplotlib, Scikit-learn, Keras, TensorFlow, Heroku, dbt, Kafka, Flask, CI/CD, ETL / ELT, Hadoop, Excel, REST API

Cloud

Amazon Web Services (RedShift, DynamoDB, Lambda, ECS, EKS, SQS, S3, Kinesis, Athena, CloudWatch, SageMaker, EMR), Google Cloud Platform (Pub/Sub, GCS, Cloud SQL, Dataproc, BigQuery, Stackdriver, GKE, GCR, GCE, GAE, AutoML)

EDUCATION

San Jose State University

MASTER OF SCIENCE IN SOFTWARE ENGINEERING

San Jose, CA

08/2017 - 12/2019

University of California, San Diego

BACHELOR OF SCIENCE IN ELECTRICAL ENGINEERING

La Jolla, CA

09/2012 - 06/2016

PROJECTS

Data Science Competitions

Takeaways: Data Analysis, Feature Engineering, ML Model Lifecycle

- Ranked **1st** in book recommender system, **2nd** in traffic image classification, **3rd** in news article text clustering, and **5th** in medical text classification in the final leaderboard of data science competitions using **Python (pandas, scikit-learn, matplotlib, scipy)**