KEVIN CHUANG
SOFTWARE ENGINEER · DATA ENGINEER

⌂ San Jose, CA ☎ (XXX) XXX-XXXX ✉ kevinchuang7@gmail.com

🖥 k-chuang.github.io github.com/k-chuang 🔗 linkedin.com/in/kevin-chuang

WORK EXPERIENCE

HomeLight *San Francisco, CA*
SENIOR DATA ENGINEER *July 2022 - Present*
DATA ENGINEER *Nov. 2020 - July 2022*

- Designed, built, and launched core algorithm to rank and match real estate agents with leads using **Python**, **SQL**, **Django**, **ElasticSearch**, **Airflow**, **AWS (Kinesis, S3, Redshift)** and **Heroku**, **boosting lead conversion rate by 15%**
- Built internal TV ad tracking, planning, & bid optimization system utilizing many disparate data sources to support marketing team and **significantly reduce agency ad costs by $250,000** per year using **Python**, **Django**, **Airflow**, **SQL**, **AWS** (**ECS**, **S3**, **RedShift**) and **Heroku**
- Architected, built, and launched data processing system to process & expose US property data (property images, property history) supporting HomeLight website pages using **Python**, **Airflow**, **SQL**, **ElasticSearch** and **AWS (Lambda, SQS, S3, Redshift, DynamoDB)**
- Developed record linkage solution of disparate data sources to support operations team in keeping track of real estate agents & transactions and discover **hundreds of thousands of dollars in missed potential revenue** using **Python, SQL, Airflow, Redshift, ElasticSearch**

ViacomCBS *San Francisco, CA*
MACHINE LEARNING ENGINEER *Feb. 2020 - Nov. 2020*
DATA SCIENCE INTERN *June 2019 - Dec. 2019*

- Designed, built, and launched ML system consisting of web service & event-driven batch processing jobs to automatically generate intelligent video previews using **Python**, **Django**, **Docker**, **Kubernetes** / **Helm**, **GCP** (**Stackdriver**, **GKE**, **Pub/Sub**, **GCS**, **Cloud SQL**), **Codefresh** and **Grafana** to save company money from using external services and support internal brands (**CBS News**, **CBS All Access**)
- Developed serverless, batch processing collaborative filtering recommendation system using **Python**, **Spark**, **Redis**, **Kubernetes**, and **GCP** (**BigQuery**, **Dataproc**, **GKE**) to power personalized movie / show recommendations for **CBS All Access**
- 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 *Mountain View, CA*
DATA ENGINEER *Feb. 2017 - May 2019*
TEST ENGINEERING INTERN *Oct. 2016 - Feb. 2017*

- Designed & implemented **Python** data pipelines for verification, processing, and transfer of large amounts of critical audio data into **MongoDB** database to enable cross-functional engineering teams

SKILLS

Programming	Python, Bash, Golang, JavaScript, MATLAB
Database	PostgresSQL, 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
Cloud Technology	Amazon Web Services (RedShift, DynamoDB, Lambda, ECS, EKS, SQS, S3, Kinesis, Athena, CloudWatch, SageMaker), Google Cloud Platform (Pub/Sub, GCS, Cloud SQL, Dataproc, BigQuery, Stackdriver, GKE, GCR, GCE, GAE, AutoML)

EDUCATION

San Jose State University *San Jose, CA*
MASTER OF SCIENCE IN SOFTWARE ENGINEERING *Aug. 2017 - Dec. 2019*

University of California, San Diego *La Jolla, CA*
BACHELOR OF SCIENCE IN ELECTRICAL ENGINEERING *Sept. 2012 - June 2016*

PROJECTS

Yet Another URL Shortener *Takeaways: Microservices, AKF Scale Cube, System Design*

- Collaborated with team in hackathon to design, implement, and launch URL shortener web service that satisfies **AKF Scale Cube** with **React.js** frontend, **Kong** API Gateway, multiple load balancers, **Golang** API microservices, and sharded **MongoDB** clusters hosted on **AWS**

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)**

Utopia Alexa Skill *Takeaways: REST API, Web Scraping, Software Testing, Code Coverage*

- Developed and deployed a serverless Alexa skill to help people with depression using **Python** and **AWS (Alexa Skills Kit, Lambda)**