

KUANG-HSUAN LEE

(+1)858-336-1540 ◇ light11470617@gmail.com ◇ 180 Brannan St, San Francisco, CA 94107

<https://light0617.github.io/> ◇ <https://www.linkedin.com/in/khl1147/>

EDUCATION

University of California - San Diego

Sep 2016 - Mar 2018

M.S. Computer Science and Engineering

National Taiwan University

Sep 2012 - Jan 2016

B.S. Computer Science & Information Engineering: GPA 4.14/4.3 (top 5%)

M.S Finance (focus on Financial Engineering): GPA 4.15/4.3

EXPERIENCE

Ads Machine Learning Engineer – Yelp Inc., San Francisco

Aug 2019 - present

- Figured out several features and projects to improve Ads targeting model from 10 papers
- Designed a new framework for Ads Distance Recall and implemented the framework to improve RPO **5%** for restaurants and presented it in front of 50 engineers
- Migrated categorical Statistics from Hadoop to **Spark** for Ads targeting model
- Delivered Yelp Delivery feature for Ads with ElasticSearch, Java and Python
- Applied Spark to do feature verification batch for **Ads targeting model**
- Improved offline evaluation system to precisely analyze data on different conditions with Python
- Presented six recommendation system **papers** for Ads team and presents papers to improve productivity

Software Engineer – Hitachi Vantara, Santa Clara

May 2018 - Aug 2019

- Designed and developed a real-time log tracking web application to visualize 10k+ logs with Vue.js and GoLang
- Improved **60%** performance with routing sharding for Lantern
- Developed Restful APIs for security product with Java Spring Boot, ElasticSearch and Docker

PROJECT

Fake Chinese Word Generation for Art Exhibition

Nov 2019 - Jan 2020

- Collaborated with a student from School of Visual Arts to apply Variational Autoencoder Generative Adversarial Network (**VAE GAN**) to generate a new Chinese word and do words morphing

Cryptocurrency Dashboard

Sep 2018 - Oct 2018

- **Github:** github.com/Light0617/CryptocurrencyDashboard
- Created a real-time cryptocurrencies dashboard for portfolio management [MongoDB, Express, React, Redux and Node.js]
- Developed scalable ETL data pipeline for streaming cryptocurrencies data [Kafka, Kafka Streams, Cassandra]

Personalizing Bundle Recommendations on Steam

Nov 2017 - Dec 2017

- Improved to AUC 96% by applying user-centric diversity, price sensitivity model and **Bayesian Personalized Ranking (BPR) (Learning to Rank)** where there are 30,000 users and 700,000 items
- Generated the new bundle for the user with a greedy algorithm, user pricing sensitivity model and BPR model

WSDM-KKBox's Music Recommendation Challenge at Kaggle

Oct 2017 - Dec 2017

- Built a model to predict whether or not the given user listen to the given music in less one month where there are the 5,901,935 in training data with about 30,000 unique users and 30,000 music
- Achieved the top 36% ranking in Kaggle Competition by applying **Light GBM (Gradient Boosting)**

SKILLS

Machine Learning: deep learning, recommendation system, gradient boosting

Backend: Java, Python, C++, Golang