# Kailing Ding

markdingkl@gmail.com | +1 (415)509-9231 | Sunnyvale, CA | Work Authorized

---

Hi, I'm a data scientist with full-stack and model deployment experiences.

## EDUCATION

### UC SAN DIEGO
BS in Data Science
2018 - 2021
GPA: 3.6
Provost Honors: 2018 - Present

## LINKS

Portfolio:// **markding.me**
LinkedIn:// **kailingding**
Github:// **@kailingding**

## COURSEWORK

### UNDERGRADUATE
Data Mining
Machine Learning
Neural Networks
SQL/NoSQL Databases
Data Visualization
Natural Language Processing
Advanced Data Structures and Algo
Object Oriented Algorithms and Design
Software Engineering for IoT Apps

## SKILLS

### PROGRAMMING LANGUAGE
• Python • Java • JavaScript • SQL
• HTML • CSS

### ML KNOWLEDGE
• Time-series Forecasting • Data Visualization • Deep Learning
• Supervised & Unsupervised Learning
• Tree-based Models

### FRAMEWORKS AND LIBRARIES
• Sklearn • Pytorch • Pandas
• Numpy • ReactJs • React-Redux
• NodeJs • ExpressJs • MongoDB
• PostgreSQL • AWS EC2/ROUTE 53
• Highcharts • D3

### INDUSTRY SKILLS
• Git • Agile Development
• Unit Testing • Trello

## EXPERIENCE

### INCOMING DATA SCIENCE INTERN | EDF Energy
June 2020 – Sep 2020 | DS Team | San Diego, CA
- Will be developing an automated data processing pipeline as well as prediction models for wind resource assessment.

### FULL-STACK DATA SCIENTIST | UCSD Triple C
Jan 2020 – Present | self-employed | San Diego, CA
- Lead a scrum team and develop a sales forecasting and analytics product for restaurants that eliminate food waste. (Dashboard-liked Web App)
- Implement an automated time-series forecasting ETL pipeline for product's sales forecasting feature using Python and MongoDB.
- Utilize ReactJs for frontend dashboard as well as charting libraries such as Highcharts and D3 for interactive charts in the dashboard.

### DATA SCIENCE INTERN | GAC R&D Center Silicon Valley
June 2019 – Sep 2019 | AI Team | Sunnyvale, CA
- Optimized the prediction model for the health of electric cars' battery in production with an 8 % increase in metric by engineering features using Sklearn.
- Created time-series data visualization based on the vehicle's sensor data that helped the team determine important features for modeling.
- Wrote production-quality unit tests for the ETL pipeline.

## PROJECTS

### ML FROM SCRATCH  Present
Bare bone NumPy implementation of machine learning algorithms from linear regression to tree-based models.
- Used: Numpy

### DATA SCIENCE COMPETITIONS  2020
Compete in solving real-world big data problems.
- 2nd place in UCSD DSC 190A Data Mining Challenge (2nd/125, Airbnb Data)
- 2nd place in UCSD Datahacks 2020 (2nd/213, Uber Data)

### TIKTOK'S DATA VISUALIZER  March 2020
An interactive data visualization for TikTok's growth statistics using multiple JS charting library.
- Used: Highcharts, Responsive Design, TikTok's API, HTML, CSS, Github Page

### MASTER OF CLASS  Sep 2019 – Jan 2020
Built a website with authentication that allows UCSD student to search their courses' group chat QR code. (1000 users)
- Used: ReactJs, ExpressJs, AWS EC2/Route 53, JWT, Bcrypt, React-Select, Redux, HTML, CSS

### PASSWORD STRENGTH CLASSIFIER  June 2019
Implemented Gaussian Naive Baye's Classifier which is able to classify a password's security level.
- Used: Sklearn, Pandas, Numpy, Colab Notebook

## INVOLVEMENT

- Student Tech Lead at UCSD Triple C (A Technology-driven Business Incubator)
- VP of Technology at Eta Omega Chi (Professional Fraternity)
- Adventurous Sports (Scuba Diving, Snowboarding, etc.)