

**Objective:** Highly motivated and detail-oriented Data Scientist with a strong foundation in data analysis and machine learning. Skilled in data preprocessing, feature engineering, and model evaluation. Proficient in popular tools and libraries such as TensorFlow, scikit-learn, and Tableau. Experienced in working with diverse datasets and implementing analytics solutions. Committed to leveraging data-driven approaches to gain valuable insights and contribute to business growth. Passionate about continuous learning and applying new methodologies to solve complex problems.

---

## **Education**

Master of Science in Applied Data Science, University of Southern California, 2024

Bachelor of Science in Statistics, University of California Los Angeles, 2018

## **EXPERIENCE**

**American Excel Enterprises Inc** - Newport Beach, CA

**Oct 2020 - Oct 2021**

*Marketing Analyst*

- Crafting and evaluating methods for collecting data, such as customer surveys, questionnaires. Using statistical software to analyze data related to a particular product or service
- Assisting in the development of large-scale data structures and pipelines to organize, collecting and standardizing data that helps generate insights and addresses reporting needs
- Converting complex data into understandable graphs, tables, and written monthly product reports. Make sales report, the system feedback report weekly
- Correct the error in system, summarize the reason, give suggestion to software analyst for upgrading the data system. Using data for making a recommendation report to each customer's product.
- Collaborate with engineers, product managers, and data scientists to understand data needs, representing key data insights in a meaningful way. Collaborate with engineers and customers to ensure deliverables and requirements are aligned.

**Wanhe Group - Remote**

**Dec 2019 - Nov 2020**

*Business Intelligence Engineer in E-commerce*

## **PROJECT/ PROGRAM**

**Health Care Data Management Project** - LA, CA

**Aug 2021 - Dec 2021**

*Machine Learning Algo, Data analyst and UI Design, frontend*

- Use python(pandas, numpy, matplotlib pyplot, statsmodels.api, sklearn,seaborn) to clean/transform the data. Build a connection between data management and machine learning. Allow users to explore data in the database
- Use a cloud database to store text data(AWS RDS) and image data(AWS S3). Use Amazon EMR to run Pyspark Mllib for training the model, saving the model and their features in AWS S3.Connecting with Apache Kafka to read the continuous streamed data into database(json,API end point))
- Use Spark for feature extraction by developing algo in Spark SQL/Data frame.Utilize Spark MLib/Scikit-Learn for training ML models in parallel. For text data, use with Pyspack to train Logistic regression, Naive Bayes, Random Forest. For image data, use Convolutional Neural Network backed by Tensorflow& Keras.

**Stocks Investment Web App** - LA, CA

**Jan 2022 - May 2022**

*Project Manager and Coordinator; Data Analyst, Date Preprocessing, Machine Learning Algorithms*

- Utilized yfinance to extract financial data from Yahoo!Finance; Leveraged Python for data cleaning, transformation, analysis, and visualization
- Employed quantstats for quantitative finance-oriented analysis of individual stocks and portfolios; Developed optimized portfolios using PyPortfolioOpt based on user-defined constraints, objectives, and risk preferences; Utilized MongoDB to store and manage financial data.
- Built a restful API using Python Flask, enabling front-end users to exchange data with the back-end. Utilized Vue.js as the front-end framework; Implemented user authentication using JWT mechanism

- Developed machine learning models (including linear regression, natural language processing, ARIMA, GARCH, and LSTM) to predict the closing stock price of different companies

## **Yelp Review Website& IOS App - LA, CA**

**Aug 2022 - Jan 2023**

### *backend& frontend and UI design*

- Creating backend applications using JavaScript/Node.js on the server side with Express framework.
- Creating Server-Side Scripting using the Python programming language and Flask framework.
- Using with the Google API, Yelp fusion API, Using use the ipinfo.io API to fetch the geolocation if the location checkbox is checked
- Creating web pages using HTML, CSS, JavaScript, DOM, JSON format and XMLHttpRequest object.
- Implementing backend( .py; .yaml; .txt) in the cloud on Google Cloud App Engine, AWS or Azure.
- Using Angular, TypeScript, and Bootstrap on the client side and creating responsive front-end.
- Using HttpClientModule of Angular for AJAX.
- Angular framework& node.js: <https://my-csci-hw8.wl.r.appspot.com>
- IOS App: [https://drive.google.com/file/d/1gZ\\_roQyhHSiPVLqJou94ShsW-vOTuDQF/view?usp=sharing](https://drive.google.com/file/d/1gZ_roQyhHSiPVLqJou94ShsW-vOTuDQF/view?usp=sharing)

## **Recommendation System Project**

**Jan 2023 - May 2023**

### *Data Developer, data mining*

- Conduct exploratory data analysis, develop feature data pipelines, and build predictive models using machine learning techniques
- Use machine learning techniques and algorithms, particularly in recommendation systems (collaborative filtering, content-based filtering, hybrid filtering)
- Continuously monitor and improve the performance of recommendation systems

## **SKILLS**

**Computer:** Object-Oriented Programming(R, Python), **Databases** (Google Firebase, Amazon DynamoDB, MongoDB, SQL,MySQL,NoSQL) **Software stacks** (AWS EC2, Apache Hadoop, Spark) **Data Storage**( Google Cloud Storage, Amazon S3&RDS, GCP), HTML, CSS, JavaScript, DOM, JSON, XMLHttpRequest, Angular, TypeScript, Bootstrap, AJAX, Frameworks(Flask, React, Node.js), **Web Services** (REST), Website Design (Bootstrap), Tableau, Github, Github Pages, Data Mining(Pyspark, Hadoop, Recommendation System, Spark)