Thomas Kim

714-833-0362 | thomasekim10@gmail.com

EDUCATION

University of California, Berkeley | Berkeley, CA **August 2019 - May 2023**

- **B.S.** in **Electrical Engineering and Computer Sciences**
- **Minor** in **Data Science**
- Relevant Coursework: Data Structures, Algorithms, Artificial Intelligence, Discrete Mathematics /Probability Theory, Optimization Models, Machine Structures, Designing Information Systems, Multivariable Calculus, Principles of Data Science, Databases

SKILLS

Languages/Tools: Java, Python, Flask, SQL, Linux, Pandas, Scikit-learn, HTML, CSS, Jira, Tableau, Statistics

Data Structures and Algorithms

- Object-oriented programming and functional programming
- Advanced data structures, search algorithms, asymptotic analysis and time/space complexity analysis

Git Version Control, SSH, Agile Methodology

PROJECTS

Tic-Tac-Toe AI | *Python* **August 2023**

- Constructed a Tic-Tac-Toe web application capable of two-player local and single-player AI-based gameplay using Flask-SocketIO
- Integrated Monte Carlo Tree Search based on UCT formula for AI strategy to generate an optimal move given the current game state
- Extended application functionality to allow users the option between a 3x3, 4x4, or 5x5 board size
- Designed intuitive and user friendly UI using HTML/CSS and deployed app on tttmontecarlo.pythonanywhere.com

NBA Draft Analysis | *Python* **June 2023**

- Implemented Python program using Beautiful Soup to pull data from HTML object to create a 1,000-line CSV file
- Aggregated and cleansed 20 years of NBA/NCAA basketball data with SQL into Pandas DataFrame and determined important features
- Trained a Machine Learning (ML) model with decision tree classification to predict NBA draft position for each college player
- Drafted and published an article containing embedded code blocks and visualizations using Matplotlib from DeepNote notebook

Walmart Sales Forecasting | *SQL* **July 2023**

- Developed custom SQL queries to view, aggregate, and perform data cleansing on Kaggle dataset with over 150,000 entries
- Created four unique Machine Learning Models to generate sales predictions with an accuracy greater than **95%** on test set
- Enhanced performance by using ensemble averaging that outperformed the root mean squared error of the top performing model by **6%**
- Leveraged Tableau to create meaningful visualizations about yearly performance and future sales forecast

MOOCbase | *Java* **May 2022**

- Implemented Java-based SQL Relational Database Management System (RDBMS)
- Added support for join algorithms and query optimization
- Included multi granularity locking to support concurrent execution of transactions
- Implemented write-ahead logging and support for savepoints, rollbacks, and ACID compliant restart recovery

PROFESSIONAL EXPERIENCE

Strolleta Inc. | Software Engineer Volunteer **July 2021 - September 2021**

- Developed voice cloning app built with Python/Django where users record voice samples and generate text-to-speech in their own voice
- Improved existing voice clone program to effectively function for sentences longer than 100 words
- Utilized AWS Simple Storage Service (S3) bucket into heroku application to work around storage restraints
- Rewrote HTML/CSS to improve UI and make voice clone application more user friendly

LEADERSHIP AND EXTRACURRICULARS

Data Science Society | *Jupyter/Python* **August 2020 - December 2020**

- Dedicated 10hrs/week to explore and process financial market datasets to generate a semester-long research project
- Collaborated with a diverse team of 6 to create a 25-slide final presentation deliverable for a bi-annual Symposium
- Utilized Jupyter Notebook to efficiently combine technical data to determine the performance of 10 NASDAQ stocks