JACQUELINE LEE

jacquelinekclee@yahoo.com | linkedin.com/in/jacqueline-kc-lee | github.com/jacquelinekclee

EDUCATION

University of California San Diego Expected: Mar 2023

- B.S., Data Science; Economics minor; Business minor
- **GPA**: 3.97 | Provost's Honors 10/10 Quarters
- **Courses**: Practice and Application of Data Science; Intro to Machine Learning; Business Analytics; SQL – Intro to Data Management; JavaScript/D3 – Intro to Data Visualization; Deep Learning

Skills: Python (Pandas, Sci-kit, PySpark), SQL (Postgres, Hive), Git, AWS Certified Cloud Practitioner

WORK EXPERIENCE

***Software Engineer Intern* – BallerTV** (Series B Startup) Remote | Jun 2022-Aug 2022

- Developed a program that cleans basketball scoring data, extracts relevant statistics, and feeds them into Natural Language Generation tool to automatically create customized headlines for thousands of games.
- Created dashboards (Redash, SQL) that highlight relevant metrics regarding travel costs and asset tracking to enable product managers and stakeholders to make cost-saving decisions.
- Programmed a workflow that automatically sends text messages to hired contractors when an event is postponed by leveraging Active Record and Twilio API, ensuring timely and appropriate communication.

***Fellow* – Bessemer Venture Partners** San Francisco, CA | Jun 2022-Aug 2022

- 1 of 20 fellows chosen out of 2,000+ applications and 60+ finalists; portfolio company directly reached out.
- Studied entrepreneurship with founders and product leaders in the venture capital space.

***Data Science Intern* – TE Connectivity** Remote | Jun 2021-Jun 2022

- Analyzed deep learning forecast model results with Pandas and identified 1,000+ highly accurate parts making $300M+ in revenue to educate executives on cost-saving forecast model and promote its adoption.
- Developed a fuzzy matching algorithm in Python that maps predictions to actual sales, quantifies success of marketing campaigns, increased accuracy by 50%, and was adopted on 2 projects used by 5+ departments.
- Coded statistical methodology using Python/SciPy that randomly splits potential customers into a test and control groups, reducing manual labor and enabling marketers to accurately evaluate their campaigns.

***Data Science Tutor* – Halıcıoğlu Data Science Institute** San Diego, CA | Oct 2020-Jun 2022

- Enhanced students' Python and Java skills in topics like object-oriented programming, recursion, and data structures by helping them understand assignments and coding logic during 4 weekly office hours.
- Reinforced lecture material through creating weekly quizzes, moderated the question and answer forum to resolve students' issues, and contributed to the creation of programming assignments.

PROJECTS

Random Forest, XGBoost Classifiers (Pandas, Scikit-learn) – Predicting NBA Players' Position

- Trained 4 models and gathered 30+ years of NBA data to explore whether statistics are meaningful in determining a player's position and see if the definitions of positions have held up over time.
- Increased accuracy by 30 percentage points (~40% to ~70%) with feature engineering and feature selection.
- Found model accuracy decreases over time and certain players get misclassified consistently, indicating how basketball is evolving into a "positionless" sport.

Web Scraping and Natural Language Processing (Beautiful Soup/SciPy) – Has Hip-Hop Gotten Worse?

- Analyzed the lyrics of the best rap songs from 1990-2020 to see if rap music has qualitatively declined.
- Determined the proportion of unique rhymes to all rhymes was the strongest, most significant metric by using p-values; but, found no sound evidence indicating a decline in lyrical quality of rap over time.