

EDUCATION

The University of California, Berkeley, Class of 2023

B.A., Data Science with a Concentration in Business Analytics

San Jose State University, San Jose, Class of 2025

Master of Engineering with Concentration in Data Science and Statistics

Relevant Courses: Data Structures and Algorithms, Computer Programming, Probability for Data Science, Data Inference and Decisions

PROFESSIONAL EXPERIENCE

Data Science Intern | Loxz Digital | Berkeley, CA

Jan 2023 - Jul 2023

- Designed information systems and hosted a full stack web page using Python, HTML, AWS, and CSS responsible for demonstrating marketing models with documentation on product features and improving market outreach by **27%**
- Implemented text data mining techniques for sentiment analysis on email metadata, providing actionable insights to help train and deploy a data driven machine learning model using Tensorflow, Pytorch, and Streamlit for email service clients to filter spam, resulting in a **25%** reduction in fraudulent email transactions in the first two months
- Led the design and implementation of a comprehensive ETL pipeline to integrate email marketing data from various sources and load it into a centralized data warehouse, resulting in a **31%** improvement in data processing efficiency
- Collaborated with a cross-functional team to establish robust data architectures that met business requirements and develop a natural language processing (NLP) algorithm for a telecommunication marketing client, resulting in a **30%** increase in click-through rates and conversion rates
- Automated end-to-end reporting process utilizing OpenAI API and built data visualization report in Tableau
- Updated outdated model and optimized by reducing bottlenecks by **47%** and saving the team **20+** work hours
- Conducted data mining to identify patterns in customer behavior, leading to a **15%** increase in engagement
- Implemented and managed data center using AWS cloud computing services, optimizing cloud infrastructure to deploy machine learning models for increased reliability, scalability, and security

PROJECTS

Instacart Data Analysis | EDA Project

- Performed data analysis on Instacart data comparing the business operations, capacity planning, and KPIs with competitors, generating information that could influence revenue by **13%** and reduce operational costs by **27%**

Email Spam Classification Model | NLP Project

- Built a NLP classification model with **95%** accuracy on JupyterHub to identify and predict spam emails using feature engineering, cross-validation, Numpy, and Pandas and visualized data with Seaborn and Matplotlib

NBA Team Analysis Model | MACHINE LEARNING Project

- Led a team and built statistical machine learning models (KNN, Random Forest, OLS regression) using Matplotlib, Seaborn, Numpy, and Pandas that predicted with **93%** accuracy the 2023 NBA teams in the playoffs

Cook County Housing Prediction Model | Predictive Data Analytics Project

- Implemented a predictive analytics model using Scikit-learn to predict housing prices based on different features

Rotaract Maps Web Page | COMPUTER SCIENCE Django Stack Project

- Created a responsive, interactive web page to help map Rotaract clubs globally and connect to broaden our organizational impact and outreach utilizing Django, SQL, and Python
- Managed the database for 2-3 years, continuously implementing a custom map generation feature with embeddings

TECHNICAL SKILLS

Skills: Java, Python, Pandas, HTML, C++, CSS, SQL, JupyterHub, AWS, Excel, Git, Tableau, PowerBI, Scikit-learn, TensorFlow, PyTorch, JavaScript, MongoDB, Express.js, React, Node.js, XGBoost, Apache Spark, Apache Hadoop

