

Karina Chen

650-861-9286 | karina.e.chen@gmail.com | [linkedin.com/in/karina-e-chen](https://www.linkedin.com/in/karina-e-chen) | github.com/karinaeichen

EDUCATION

University of California, San Diego

Sept. 2020 – Mar. 2024

Data Science B.S., Minor in Design

La Jolla, CA

- GPA: 3.98
- Relevant coursework: Calculus & Analytical Geometry, Data Analysis & Inference in R, Data Management, Data Structures & Algorithms, Data Visualization, Deep Learning, Image Processing, Interaction Design, Linear Algebra, Machine Learning, Mathematical Statistics, Probabilistic Modeling and ML, Recommender Systems & Web Mining, Representation Learning, Systems for Scalable Analytics
- Honors: Magna Cum Laude, Provost Honors (every quarter)

EXPERIENCE

American Express

Sept. 2024 – Present

Software Engineer

Phoenix, AZ

- Driving Global Compliance Intelligence platform modernization efforts within the data engineering team, utilizing Google Cloud Platform (GCP), Python, and SQL to enhance data consolidation workflows
- Designing and developing an end-to-end test automation framework to streamline validation of data workflows by orchestrating test cases via Airflow DAGs, with tables managed in BigQuery
- Supporting transactional data processes by enhancing Airflow DAGs to conditionally execute the end-to-end test automation framework, optimizing SQL queries, and implemented table prefixing to ensure isolated, production-safe test runs
- Improving DAG performance and readability by replacing Dataproc and PySpark with BigQuery client, reducing processing overhead and simplifying data workflows

American Express

Jun. 2023 – Aug. 2023

Software Engineer Intern

Phoenix, AZ

- Led planning and developed enhancements for an internal Financial Strategy Tool on the Financial Data Engineering scrum team following the Agile methodology
- Designed and implemented application home page featuring adaptive navigation to internal pages and external applications using React.js, HTML, and CSS
- Identified and corrected data discrepancies from backend JSON output and configured dynamic page layout and tooltips for displaying information on data dashboard
- Built custom AI chatbot using RASA framework that can answer 30+ application FAQs to improve usage efficiency and domain understanding

PROJECTS

Diffusion Models for Image Generation | *Python, Pytorch, Git*

Sept. 2023 – Mar. 2024

- Implemented diffusion models for data science capstone project with mentors Alex Cloninger and Rayan Saab
- Clarified the diffusion process using a spiral distribution in Pytorch (github)
- Investigated how the 3D properties of depth, saliency, and illumination are encoded within the internal activations of latent diffusion models using linear probing (website)

Analysis and Prediction of Survivor Winners (link) | *Python, pandas, numpy, scikit-learn*

Aug. 2023

- Converted survivoR data package from R to Python datasets, cleaned data, and restructured data frames by merging and dropping columns in order to analyze all 42 seasons and 626 contestants of the show Survivor
- Used pandas, numpy, scikit-learn, matplotlib, and seaborn to clean dataframes, perform feature engineering, and apply models to predict Survivor season winners

TECHNICAL SKILLS

Languages: Python, SQL, Java, JavaScript, HTML/CSS, R, LaTeX

Developer Tools: Git, Google Cloud Platform (BigQuery, Airflow), VS Code, Jupyter Notebook, Microsoft Office, Rally

Frameworks: pytest, React

Libraries: pandas, NumPy, Matplotlib, scikit-learn, D3

Certifications: Google Cloud Professional Data Engineer (Jun 2024)