

Meilin (Nicole) Pan

meilinp@usc.edu | 805-689-5974 | linkedin.com/in/meilin-pan-8886a6251/ | <https://github.com/MeilinP> | Los Angeles, CA

SKILLS

- **Programming:** Python (**NumPy, Pandas, Scikit-learn, PyTorch, TensorFlow**), SQL, R, SAS, Big Data (Spark, Hive), R, SQL, Cypher, MATLAB
- **Technologies:** AWS EC2, Google Cloud Platform (Cloud Computing), Tableau, Matplotlib, Seaborn (Visualization), MongoDB, DynamoDB (Database management), SAS, SPSS, Excel (Statistical tools)

EXPERIENCE

Namefi.io

Palo Alto, CA

Data Analyst Intern

Jul 2023 – Dec 2023

- Partnered directly with the founder as the first data hire to formulate and report on Ethereum Name Service (ENS) KPIs
- Built an automated weekly dashboard using SQL and Dune Analytics to visualize ENS market trends and received recognition from the founder of Dune on Twitter
- Employed time series analysis techniques to forecast domain name registration growth and renewal rate, improving accuracy by 15% through integrating various databases
- Collaborated with product manager to perform cohort analysis to identify potential new customers and lead to success in obtaining 500+ new community members on Discord in 2023

Baidu, Inc.

Beijing, China

Data Analyst Intern, Key Account Management Department

May 2021 – Aug 2021

- Developed CPC (Cost Per Click) pricing model to generate 15% higher revenue through optimized pricing recommendations for different advertisement display positions, aligning advertisement strategies with client goals
- Leveraged user data to produce detailed Search Engine Marketing (SEM) campaign performance reports for top Tech-driven Retail companies (Meituan, Didi Global, Ctrip), saving 90 hours of monthly labor
- Managed and analyzed extensive daily marketing performance data encompassing 400K+ records using advanced SQL queries and Excel VLOOKUP function, facilitated by cloud-based storage solutions

PROJECT

Audio Event Detection & Notification System (Android & AWS)

Dec 2024

- Developed an automated audio event detection pipeline using Termux (Android), AWS EC2, and SageMaker to classify environmental sounds and trigger email alerts.
- Designed an AWS-based audio processing workflow leveraging Librosa for feature extraction and an AWS SageMaker endpoint for classification (e.g., detecting vehicles, gunshots, or engine sounds).
- Automated data flow with **cron jobs** and **Systemd timers**, ensuring continuous recording, transfer, and classification.
- Integrated SMTP-based email alerts to notify users when specific sound events exceed probability thresholds.
- Insured secure AWS authentication and email handling using .env files for credentials and environment variables.

Home Credit - Credit Risk Model Stability

May 2024

- Conducted extensive data wrangling on 32 CSV and Parquet files containing 465 variables using Pandas and Polars.
- Developed machine learning models to predict loan default risks, employing algorithms from scikit-learn and Keras
- Improved loan accessibility for individuals with limited credit history

E-Commerce Database Management System

May 2024

- Developed a Flask-based RESTful API with SQLAlchemy ORM, handling CRUD operations for users, products, and orders
- Designed a PostgreSQL relational database, enforcing foreign key constraints and indexing for performance
- Built an admin dashboard for user, product, and order management with real-time updates
- Utilized MySQL on AWS EC2 to implement a relational database and leveraged distributed database for efficient data management

EDUCATION

University of Southern California

Los Angeles, CA

Master of Science in Applied Data Science

Jan 2024 – Dec 2025 (Available for full-time from May 2025)

Coursework: Machine Learning, Data Science Statistics, Big Data Management

University of California, Santa Barbara

Goleta, CA

Bachelor of Science in Statistics and Data Science GPA: 3.57 / 4.0 (Dean's Honors)

Sep 2019 – Mar 2023