Felipe Lorenzi

LorenziFelipe7@gmail.com | linkedin.com/in/felipe-lorenzi | (914) 844-9940 | San Diego, CA

WORK EXPERIENCE

Graduate Student Researcher   Apr. 2024 – Sep. 2025
UC San Diego   *San Diego, CA*

- Conducted research investigating robust and trustworthy data analytics under Dr. Babak Salimi.
- Co-authored a paper submitted to VLDB 2025 introducing a framework for simulating data quality issues and exposing crucial vulnerabilities often overlooked in even the most robust ML pipelines.
- Led two research projects on conformal prediction, a robust uncertainty quantification method for AI/ML models.

Data Analyst   Dec. 2022 – Oct. 2023
Hagerty Consulting   *San Diego, CA*

- Built tools and processes from scratch in Python, Excel, and VBA for regulatory compliance in an improvised, emergency response setting.
- Mediated communications with stakeholders and cross-functional teams belonging to four different organizations.
- Enhanced data collection processes for improved availability and accuracy while achieving speed ups of over 60%.
- Conducted QA/QC and developed remediation plans to reduce data discrepancy rates from over 10% to less than 5%.

Data Scientist Intern   Apr. 2022 – Jul. 2022
Neighborly   *Waco, TX*

- Utilized SQL, statistics, and machine learning models in Python to model franchise success rates in a self-led analysis.
- Presented analytical findings to C-suite executives and collaborated to conceive actionable insights.
- Streamlined survey design and data collection pipeline for efficiency and compatibility with future company datasets.
- Managed project timelines independently while working closely with the business intelligence team.

Intern   Jun. 2020 – Oct. 2020
Catavento Book Distributor   *São Paulo, Brazil*

- Trained and deployed a machine learning model in Python to rectify genre misclassifications of more than 30,000 books by leveraging natural language processing transformer models.

Projects

- **Data Management:** Collaborated in a team of three to create recommender system trained on tens of millions of data points scraped from Goodreads using Python, Docker, PostgreSQL, Neo4j, and Qdrant. (2025)
- **Forecasting:** Built database of daily car sales from Cars.com by creating web scraper in Python and analyzed for trends, saving money when purchasing a car a few months later. (2023)
- **Machine Learning:** Analyzed brain EEG signals when listening to music using signal processing and machine learning in Python using libraries such as mne, scipy, librosa, and pytorch. (2022)
- **Statistical Inference:** Analyzed dataset of NYPD civilian complaints using Python to uncover evidence suggesting systematic bias against certain minority groups. (2021)

EDUCATION

University of California, San Diego   April 2026 (expected)
MS, Data Science   *San Diego, CA*

- Graduate research in robust and trustworthy data analytics (uncertainty quantification for AI).

University of California, San Diego   August 2022
BS, Cognitive Science with Specialization in Machine Learning   *San Diego, CA*

- Data Science minor. 3.7 GPA