Eduardo Escoto

Contact

Learn more at: eduardo-escoto.com
Connect with me at: linkedin.com/in/eduardo-escoto
See what I'm working on at: github.com/eduardo-escoto

Talk to me at: +1 (858) 357-1868
Write me at: e1escoto@ucsd.edu
Find me at: San Diego, CA

Objective

I am a second-year Ph.D. student funded by the National Science Foundation, working under the guidance of Dr. Julian McAuley. My current research interests are in Representation Learning (Autoencoders, Vector Quantization) and Generative Models(Diffusion, Flow Models), and their applications to Music and Recommender Systems. I am also interested in the concept of model memory for LLMs and Transformer-like models.

Honors and Awards

- **NSF CSGrad4US Fellowship** — Awarded: August 2022
  National Science Foundation
  - **Three Years of Full Funding**: The National Science Foundation is funding my Ph.D. in Computer Science at UC San Diego for three out of five years.

Education

- **University of California, San Diego** — San Diego, CA
  Doctor of Philosophy in Computer Science — September 2024 – June 2029 (Expected)
  - **Advisor**: Dr. Julian McAuley
  - **Research**: Recommender Systems, Generative Audio and Music
  - **Concentration**: Machine Learning and Artificial Intelligence.
- **University of California, Santa Barbara** — Santa Barbara, CA
  Bachelor of Science in Mathematics and Bachelor of Science in Statistics — September 2017 – June 2020
  - **Concentration**: Probability and Statistics.
  - **Organizations**: Data Science Club, Association for Women in Mathematics.
  - **Relevant Coursework**: Bayesian Analysis, Data Science, Machine Learning, Stochastic Systems, Real Analysis, Numerical Analysis, Complex Analysis, Graph Theory, Advanced/Abstract Linear Algebra.
- **San Diego Miramar College** — San Diego, CA
  Associate of Arts in Mathematics — September 2014 – June 2017
  - **Organizations**: Phi Theta Kappa Honors Society.
  - **Relevant Coursework**: Calculus I/II/III, Discrete Mathematics, C++, Java, Visual Basic.
  - **Accomplishments**: Honors Student, Deans List Scholar, Annual Research Symposium Presenter.

Teaching Experience

- **UCSB INT 5: Data Science Foundations** — Santa Barbara, CA
  Undergraduate Learning Assistant — September 2019 – December 2019
- **UCSB INT 15: Data Science Tools and Techniques** — Santa Barbara, CA
  Undergraduate Learning Assistant — March 2019 – June 2019

Professional Experience

- **Instacart** — San Francisco, CA
  Machine Learning Engineer, PhD Intern — June 2025 – September 2025
  - **Generative Recommendations and Catalog Applications**: I am working on implementing a comprehensive Semantic ID system to improve various systems across Instacart's infrastructure. This project aims to explore Semantic ID applications at a large scale for catalog applications, generative recommenders, and semantic deduplication.
- **University of California, San Diego** — San Diego, CA
  Graduate Student Researcher — September 2024 – Present

- **Toyota Industry Collaboration**: I am working on a project funded by Toyota in regards to leveraging memory to improve driver assistance systems and automated driving systems.

- **Noom** New York, NY
  Senior Data Scientist January 2022 – September 2024
  - **Cross-Functional Partnership**: Coordinates collaboration between Data Science, Data Engineering, and Engineering teams to ensure successful implementation of data requirements and infrastructure for event-based data consumption and model deployment.
  - **Project Leadership**: Leads a team of Data Scientists and Analytics Engineers for implementation of novel data products and features, and refactoring of existing code.
  - **A/B Testing and Experiment Design**: Aids in design and development of A/B Tests and Experiments for Growth. Analyzes experiments results and helps stakeholders in pinpointing findings and high value results.
  - **ETL and BI Development**: Develops Spark and Redshift ETL pipelines for data analysis and for use across the organization. Communicates and documents data requirements from stakeholders and other Data Scientists. Builds LookML for Looker Dashboards and Explores to use new data sources.
  - **User Behavior Analysis**: Incorporates techniques like Causal Inference, Clustering, and Machine Learning to utilize observational and experimental data in order to understand user behavior and to inform product decisions.
  - **Machine Learning**: Builds ML models for Real-Time Personalization, Recommendation, and NLP from free text entry data. Leverages LLM's for building AI driven products.

- **Afiniti** Washington, DC
  Data Scientist August 2020 – December 2021
  - **Built Traditional and Deep Learning Models**: Prototyped, engineered, and pipelined a plethora of NLP models, embedding models, and predictive models. Utilized traditional ML methods like Gradient Boosted Decision Trees and Random Forests for models with tabular data.
  - **Bayesian Data Analysis & Modeling**: Created STAN models to model interaction between call center agents and callers. Researched and improved models through iteration and testing.
  - **Code Contribution and Numerical Computing**: Maintained company code repositories for Machine Learning and Modeling. Implemented and optimized new features in our modeling package.

- **UC Santa Barbara Health and Wellness** Santa Barbara, CA
  Data Science Intern May 2019 – June 2020
  - **Facilitated Data Gathering and Management**: Leveraged platforms like Qualtics and Microsoft's Power BI in order to centralize data gathering and availability. Monitored survey question quality, and response quality.
  - **Analyzed Data and Builds Models**: Incorporated a wide array of tools and techniques to analyze all departmental and survey data. These included Bayesian Methods, Machine Learning, and Classical Methods.
  - **Optimized Internal Processes**: Innovated departmental systems and replaced them with modern analogues. Built web–based solution to streamline inventory and key management with a custom React-GraphQL App.

- **UC Santa Barbara Computer Science** Santa Barbara, CA
  Computer Science Undergraduate Learning Assistant March 2019 – June 2020
  - **Created Curriculum**: Worked with team to meticulously write guided labs and homeworks in Jupyter Notebooks. Created resources for students to use in minimizing programming learning curve.
  - **Engineered and Maintained Autograding Tooling**: Developed python scripts in order to autograde Jupyter Notebook assignments and upload scores and automatic feedback to Grading Portal via API.
  - **Held Office hours and Led Review Sessions**: Spent time tutoring and helping students with homeworks and labs. Created midterm and final review questions and went over them in a lecture style session.

- **Biosero Inc.** San Diego, CA
  Software Engineering Intern September 2016 – September 2017
  - **Database Engineering and Driver Development**: Engineered and optimized microplate tracking database. Developed drivers for a myriad of Bio-tech instruments for use in automization. These included Robotic Arms Sealers/Desealers, Acoustic Dispensing Systems, Microplate Readers, and many more.
  - **MagneMotion Track Integration**: Implemented magnetic track system for use in large scale and connected integrations. Utilized graph theory and queueing theory to streamline track movement algorithm.

- **Starbucks Coffee** San Diego, CA
  Shift Supervisor and Barista June 2013 – June 2019
  - **Certified Coffee and Food Master**: Succeeded in Coffee and Food master certifications and continued on to win the district Barista Championship in 2016. First Partner of the Quarter at store.
  - **Established Leadership and Teamwork Skills**: Led a team of world-class baristas to create a positive, inclusive, and motivating work environment that customers consistently enjoy visiting.

Professional Certifications

- **Deep Learning Essentials** CoRise
  Credential ID: 69989482 Issued: March 2023
  - **Foundations of Deep Learning**: Learned the foundations of deep learning and practiced with training and implementing neural networks in Pytorch while covering topics convolutional neural networks (CNNs), transformers, and generative adversarial networks (GANs).
- **Natural Language Processing** CoRise
  Credential ID: 70057964 Issued: March 2023
  - **Core NLP Building Blocks**: Developed an understanding of core NLP components — word vectors, intent classification, entity recognition and many more using transformer architectures like BERT and GPT — while building projects like embedding-based retrieval and smart-compose.
- **MLOps: From Models to Production** CoRise
  Credential ID: 67862113 Issued: February 2023
  - **Building Effective Real-World ML systems**: Learned about bootstrapping datasets, improving label quality, experimentation, model evaluation, deployment, and observability through hands-on projects. Developed familiarity with bridging the gap between State of the Art ML and real world systems.
- **Data Centric Deep Learning** CoRise
  Credential ID: 64652793 Issued: December 2022
  - **Learned Data-Centric Approach to Deep Learning**: Simulated the real world challenge of improving data quality, building and testing deep learning models, and improving performance with a human-in-the-loop in both natural language and computer vision models. Built a web application with an embedded model.
- **Building Transformer-Based Natural Language Processing Applications** NVIDIA
  Credential ID: 84adff79bea24686b8dc58f877397d71 Issued: September 2022
  - **Built and Deployed Transformer Models**: Processed text data in preparation for Transformer based models, built Transformer based models for NLP tasks, and deployed them for real time inference.
- **Personalized Recommendations** CoRise
  Credential ID: 56311061 Issued: August 2022
  - **Hands-on Learning of Recommender Systems**: Built projects that leveraged multi-stage large scale recommenders and ML based approaches to ranking.

Programming Skills

- **Languages**:
  - **Exprienced With**: Python, SQL, Julia, R, Javascript/Typescript, LaTeX, Shell Scripting.
  - **Familiar With**: Rust, C++, C#, Java, Lua.
- **Technologies and Tools**:
  - **Experienced With**: Spark, Tensorflow, Torch, Metaflow, Dash, STAN, React, Git, Looker.
  - **Familiar With**: Airflow, Shiny, Docker, D3, Tableu.