Sumadhu Rubaiyat (Genie)

srubaiyat@ucsd.edu | (858) 833-5535 | github.com/srubaiyat05 | Website: srubaiyat05.github.io

EDUCATION

University of California, San Diego Expected Graduation: June 2026
B.S. Mathematics - Computer Science, Minor in Data Science – Provost Honors
Coursework: Machine Learning, Recommender Systems, Natural Language Processing, Practical Data Science, Data Analysis/Inference, Algorithms, Data Structures, Computability Theory, Computer Organization/Architecture, Machine Language, Linear/Numerical Linear/Abstract Algebra, Probability, Statistics, Multivariable/Vector Calculus, Differential Equations, Real Analysis, Combinatorics, Graph Theory, Discrete Math, Logic, Optimization
Languages: Python, Matlab, R, Ruby, Java, C, C++, ARM, HTML, JS, ReactNative, React, Swift, SQL
Libraries: PyTorch, TensorFlow, TorchIO, OpenCV, SegmentAnything, NLTK, scikit-learn, patsy, scipy.stats, statsmodels.api, SimpleITK, Torchvision, OpenAI/GPT, NumPy, Matplotlib, Seaborn, Pandas
Other tools: XCode, Expo, Android Studio, Docker, Firebase Firestore/Storage/Hosting, GDB, LaTex, Git, JUnit, NodeJS, 3D Slicer, Blender, Conda, Labelme, Raspberry Pi, Arduino, Copilot, TorchHub, Bash

EXPERIENCE

Advanced Robotics and Controls Laboratory – *Researcher* April 2023 – June 2025
- Trained Computer Vision models to simulate the movement of tissue upon interaction by a surgical robot, identify cuts in skin, align breathing lung models, etc, for up to 95% accuracy
- Read technical research papers to apply 2D ML models to 3D medical images, up to 85% accuracy
- Researching Graph ML applications for equivariance in on 3D medical image to outperform GCNNs
- Implemented DL scripts ensuring high standards for code-readability for academic journal reviewers
- Briefed progress to 30+ graduate students, postdocs, and professor Michael Yip in weekly meetings
- Utilized Cloud and GPU resources on a remote server to work with 100+ GBs of training data

UC San Diego Department of Mathematics – *Honors Thesis* January 2024 – June 2025
- Mentored by Dr. Joshua Frisch in Group Theoretic Symbolic Dynamics Research Group
- Attended an in-person Moore Method style lecture weekly lasting 3-6 hours uninterrupted
- Collaborated with 2 students for 10+ hours weekly to solve problem sets on the research frontier
- Studied *Office Hours with a Geometric Group Theorist*, covering cayley graphs, quasi-isometries, amenability, ultra-filters, ultra-limits, random walks on groups, Grigorchuk group, etc.
- produced novel result regarding embedding free groups within the normal subgroups of the automorphism group of the full shift
- presented novel results and expository content regarding symbolic dynamics in a 20-minute talk to honors committee consisting of UCSD math faculty, receiving Honors with High Distinction

UCSD Department of Mathematics – *Directed Reading Program* August 2024 – September 2024
- Read and re-implemented results from technical research papers about GNNs and GraphSage algorithm for category classification on graph/multigraph data on protein-protein interaction, reddit and citations
- Presented findings in 20 minute talk covering background, algorithm, results and theoretical backing

UCSD Department of Computer Science and Engineering – *Tutor* August 2023 – Present
- Hosted 1000+ tutor hours to teach technical topics like Intro to ML (CSE151A), Theory of Computability (CSE105), Algorithms (CSE101) and Python (CSE6R), and prerequite knowledge like proof technique, discrete math, linear algebra, etc.
- Collaborated with 15 tutors to grade 100,000+ assignments
- Enrolled in Tutor Apprenticeship to learn explanatory techniques

Russell Lab @ UCSD CSE – *Researcher under Russell Impagliazzo* May 2025 – Present
- Conducted literature study on the smallest enclosing ball problem
- Presented a 45 minute talk about the existing literature of the smallest enclosing ball problem
- Discovered scalable algorithm, pending experiments on more capable hardware

Spatiotemporal Learning Lab @ UCSD CSE – *Researcher under Rose Yu* June 2025 – Present

- Generated experimental results for symmetry discovery algorithm for KPPF equations
- Created Deep Learning Architecture for Equation Learning, pending experimental results

Gao Lab @ UCSD CSE – *Researcher under Sicun Gao* June 2025 – Present

- Proving complexity bounds for the Monte Carlo Tree Search-based "Sample and Bound" optimization technique developed by Sicun Gao

ML Interpretability Working Group – *Reading under Sanjoy Dasgupta* July 2025 – Present

- read and presented technical research papers on ML Interpretability in 30-minute talk to 40 students
- created algorithm for converting a random forest to shallow multiway decision tree, pending generalization results and runtime analysis

PUBLICATIONS

Application of Large Language Model in Clustering Low Count Non-Gaussian User Behavior Time Series – *IEEE Big Data 2024 Conference*

Feedback-centric Optimized Time-critical Recommendation of Time Series of Promotional Rewards for User Retention – *IEEE Big Data 2025 Tutorial*

On the Automorphism Groups of Shift Spaces – *UCSD Math Honors Theses Spring 2025*

ProCut: Probabilistic Cutting Topology for Autonomous Electrocautery Tissue Dissection – *In review for Robot Automation Letter Journal 2025*

Parameter Optimization for Fine-Tuning Text-to-Image Models on Limited Hardware: Balancing Runtime and Generation Quality – *Complete and Unpublished*

Detecting Emergent Symmetry in KPPF Equations via Deep Learning – *In progress*

Scalable Deterministic Smallest Enclosing Ball Algorithm in High Dimensions – *In progress*

Decision Forests to Shallow Multiway Trees – *In progress*

Sample Complexity of Sample-and-Bound – *In progress*

PROJECTS

Sequential Climate Prediction – *Pytorch, U-Net, Transformer, ConvLSTM, DANN, AdaBoost*

- Tested ensemble methods using Transformers, ConvLSTMs and Domain Adversarial Neural Networks
- Ranked 4th in a kaggle competition with 300 students

Institutional Mobile Application Framework – *ReactNative, Firebase*

- No-code app-builder template with display options for upcoming events, news, video announcements, calendars, contacts, bell schedule, interactive map, quick links, etc with an intuitive Firebase backend

Reinforcement Learning Model Predicting Flight Delays – *Tensorflow, NumPy, SKLearn*

- Predicted arrival delays from carrier, expected departure/arrival time, origin, destination, etc.

3D Medical Image Automatic Segmentation Model – *SegmentAnything, MatPlotLib*

- Created image scroller with SAM, wherein segmenting a jaw in 3D brain MRI takes 10 seconds

RentTheRunway Recommendation System – *SKLearn, Tensorflow*

- Collaborated in a team of 4 to create a 68% accurate recommendation system for RentTheRunway

Analysis of UCSD Quarter Progress against Traffic Accidents Data –*NumPy, Seaborn, Pandas*

- Conducted hypothesis/correlation tests on SDPD traffic data categorized by UCSD quarter progress and presented findings in an annotated Jupyter notebook and slideshow

AI Game Agent – *2048, Sudoku, Gomoku*

- Created AI Game Agents for 2048, Sudoku and Gomoku