

# Alp Ata Narin

alp.ata.narin@gmail.com • +1 (408) 858-8671 • [www.linkedin.com/in/alpatanarin](https://www.linkedin.com/in/alpatanarin) • Los Gatos, CA

## EXPERIENCE

<b>Lab Researcher</b> <i>Fay Lab at the University of Rochester Computational Biology Department</i>	August 2024 - February 2025 <i>Rochester, NY</i>
<ul style="list-style-type: none"><li>Built Python workflows with PAML and Trnalign to model large-scale data and automate divergence analysis</li><li>Built benchmarking frameworks applying <b>causal inference</b> and <b>A/B testing methods</b> to optimize accuracy-latency tradeoffs, mirroring product experimentation culture</li><li>Consolidated structural datasets into a unified database, streamlining analysis for cross-team research and enabling scalable queries across <b>millions of data points</b></li><li>Translated biological stability analyses into scalable <b>statistical frameworks</b>, identifying outliers and ranking system performance in ways comparable to product experimentation pipelines</li></ul>	
<b>Product Analytics &amp; ML Intern</b> <i>Align Technologies</i>	June 2024 - October 2024 <i>Santa Clara, CA</i>
<ul style="list-style-type: none"><li>Built <b>scalable SQL + Python pipelines</b> integrating structured/ unstructured datasets (~<b>2000 doctors</b>), reducing manual analysis time by ~<b>75%</b> and enabling product/finance stakeholders to track prescribing patterns in real-time</li><li>Developed <b>NLP + ML pipelines (spaCy, scikit-learn, XGBoost, PyTorch)</b> for classifying prescribing patterns and predicting treatment decisions with SHAP explainability, achieving <b>91% accuracy</b> across <b>300k+ records</b>, enabling product managers to forecast adoption trends and guiding finance teams in resource allocation</li><li><b>Developed reusable ML components and APIs in Python</b> for training, classification, and explainability, establishing standardized workflows that enabled reproducible deployment and cross-team experimentation</li></ul>	
<b>Lab Researcher</b> <i>Thornton Lab at the URMC Neurological Department</i>	August 2023 - May 2024 <i>Rochester, NY</i>
<ul style="list-style-type: none"><li>Conducted PCR, gel electrophoresis, DNA extraction, and prepared samples for Oxford Nanopore (Guppy), PacBio, and Sanger sequencing, ensuring high-quality molecular workflows for downstream genomic analysis and validation</li><li>Developed automated CI-style genomic analysis pipelines in <b>Bash/nf-core</b> with reproducibility, dataset versioning, and modular testing, analogous to production CI workflows</li></ul>	
<b>Computational Biology Teaching Assistant</b> <i>Hajim School of Engineering</i>	January 2024 - May 2024 <i>Rochester, NY</i>
<ul style="list-style-type: none"><li>Taught <b>30+</b> students <b>Python</b>-based bioinformatics workflows and guided weekly coding biostatistics projects</li></ul>	

## PROJECTS

<b>Fantasy Football Web App</b>	2025 (MVP completed, expanding features)
<ul style="list-style-type: none"><li><b>Developed a full stack application</b> with <b>React, Node.js</b>, and <b>Flask APIs</b> to manage <b>players, rosters and projections</b>, integrating dynamic data feeds for interactive team management</li><li>Built <b>MySQL</b> schemas, migrating to PostgreSQL for scalability; developed <b>REST endpoints</b> for roster updates/stats</li><li>Implementing <b>user authentication, component-based UI</b> and <b>state management</b> to support <b>real-time tracking</b></li><li>Implemented real-time player and score updates using <b>WebSockets and Server-Sent Events</b>, paired with a <b>cache write-through (Redis)</b> database strategy to ensure data consistency, reduce latency, and handle concurrent user sessions efficiently</li></ul>	
<b>Fraud Detection Database System</b>	Fall 2024
<ul style="list-style-type: none"><li>Built a fraud detection database system with schemas for transactions, users, behavior, card data, and ML predictions, scaling to <b>millions of rows</b> with built-in <b>role-based access controls, encryption</b>, and audit logs</li><li>Integrated Kaggle fraud datasets and built ML-driven fraud flagging with probability scores for real-time detection</li><li>Modeled administrator, data scientist, and auditor workflows to demonstrate scalability and alignment with real-world financial fraud detection systems.</li></ul>	
<b>Capture the Flag AI Agent</b>	November 2024
<ul style="list-style-type: none"><li>Developed an intelligent agent for a 2D grid-based Capture the Flag game, using <b>A*</b> search for real-time pathfinding and navigation through dynamic environments with obstacles and enemy zones, integrated <b>minimax with alpha-beta pruning</b> to simulate and counter opponent strategies in adversarial scenarios, and trained a <b>Naive Bayes classifier</b> to predict likely trap zones and ambush patterns based on historical movement data</li></ul>	

## EDUCATION

<b>University of Rochester - Double Major - Dean's List</b> <i>Double major: BS in Computer Science, BS in Computational Biology</i>	August 2021 - May 2025
<ul style="list-style-type: none"><li>Coursework: Artificial Intelligence, Advanced Algorithms, Database Systems, Data Structures and Algorithms, Machine Learning, Data Mining, Computational Statistics, Computational Systems, Computer Models, Computer Architecture, Linear Algebra, Discrete Math, Biostatistics, Computational Biology, Genetics, Biochemistry</li></ul>	

## SKILLS

<b>Data Science &amp; Analytics:</b> SQL, R, Data Mining, Statistical and Causal Modeling, A/B Testing, Product Analytics
<b>Frameworks:</b> Scikit-learn, PyTorch, XGBoost, SpaCy, Pandas, Numpy, SHAP
<b>Data Engineering and Visualization:</b> Data Pipelines (ETL), Database Merging, Tableau, Matplotlib, Seaborn, Jupyter
<b>Programming and Systems:</b> Java, C, Python (OOP, Flask), React, Node.js, Express, Bash, Linux, Git, APIs, CI/CD
<b>Domain Tools:</b> BioPython, HTSeq, nf-core, FastQC, TrimGalore, High-Performance Computing (Bluehive)