# Benjamin Lu

Citizenship: US | 978-387-5426 | benlu2267@gmail.com | LinkedIn: benjaminlu22

## Education

**University of Toronto** Toronto, ON
*BSc (Hons), Computer Science and Mathematics; GPA: 3.70 / 4.00* *Graduating May 2027*
**Relevant Coursework:** Operating Systems, Software Design, Computer Organization, Databases, Computer Networks, Data Structures, Algorithms, Machine Learning, Advanced Complexity Theory, Advanced Graph Theory, Linear Algebra

## Work Experience

**Shopify** | *Software Engineer Intern* May 2026 - Present
- Automated data cleanup via **Composer 3** metadata retention policy, enabling seamless migration from Composer 2
- Authored daily-granularity **BigQuery** spend table, replacing a monthly aggregate for day-level slot cost visibility
- Rewriting budget visualization in **Ruby on Rails** from a static monthly bar to a daily line graph; migrated spend table into Rails to eliminate repeated **BigQuery** API calls

**FactSet** | *Software Engineer Intern* May 2025 - August 2025
- Built internal **C#** API to efficiently aggregate financial tombstone information from **100+** investment banking sites
- Used **Selenium** and SkiaSharp to automate HTML capture and render **50** contextual tombstone screenshots per site
- **Designed LLM pipeline** with Microsoft Semantic Kernel to extract structured transaction information from screenshots and HTML, and implemented multiple layers of model-assessed confidence scoring for data reliability
- Deployed full-stack user interface leveraging API, allowing URL upload, Excel preview, and download functionality

**Vestmark** | *Software Engineer Intern* May 2024 – August 2024
- Migrated platform from WebAPI to REST across Gradle build files and CI tests, removing legacy dependencies
- Supported migration from **Java 8** to **Java 17**, resolving compatibility issues and extending AssertJ test coverage
- Streamlined annotation processing by removing redundant **Kotlin** processors, reducing platform build time by **10%**

**Vanta Leagues (Acquired by PlayVS)** | *Data Analyst Intern* May 2022 – August 2022
- Implemented data pipeline to migrate legacy Excel data into MySQL, enabling efficient querying and reporting
- Used **Python, scikit-learn, and Seaborn** to identify and visualize regression trends in **500+** outreach cases
- Proposed data-backed outreach strategies in weekly meetings, contributing to a **30%** improvement in signup rate

## Projects

**Game of 24 Multiplayer Platform** | *Python, python-socketio* | try it here
- Engineered real-time game state sync over WebSockets, propagating move events to all room clients concurrently
- Built **SQLite**-backed persistence layer storing per-account win/loss record and rating scores/updates across sessions
- Implemented head-to-head **Glicko-2** based rating system, modeling per-player volatility and rating deviation decay
- Architected room-scoped pub/sub event routing to isolate concurrent game sessions, eliminating cross-room event leakage and enabling multiple simultaneous matches across different gamemodes

**Operating System Kernel** | *C*
- Improved Pintos kernel code with multi-level feedback queue thread scheduling and dynamic priority recalculation
- Built virtual memory with per-process page tables,global frame table, lazy loading, and page eviction via swapping
- Extended filesystem with multilevel indexed inodes and indirect block pointers, scaling max file size (**512B→8MB**)

## Technical Skills

**Languages**: Python, C, C++, C#, Java, SQL, Bash, LaTeX, Kotlin/Groovy, HTML/CSS
**Libraries/Frameworks**: Apache Airflow, Ruby on Rails, Selenium, Pytorch, scikit-learn, NumPy, Seaborn, C++ STL
**Developer Tools**: Cursor, Claude, Codex, BigQuery, MySQL, Git, Docker, Jenkins, JetBrains, GCP, AWS, Gradle

## Teaching

**Complexity Theory**: Karp Reductions for NP Completeness, Fine grained complexity, ETH, PCP Systems
**Data Structures**: C programming, Balanced BSTs, Priority Queues, Graphs, Amortized Complexity, Hash Sets
**Discrete Mathematics**: Logic, Proof Writing Techniques, Combinatorics, Graph Theory, Probability