Brandon Lee

Brandonyli@berkeley.edu linkedin.com/in/brandon-lee-174243158
(626) 347-0499 brandonyli.github.io

Education

University of California, Berkeley (Spring Class of 2021)

- **Data Science Major, Computer Science Minor** - GPA: 3.59
- **Coursework**: Linear Algebra, Multivariable Calculus, Discrete Math, Probability Theory, Computer Science, Data Structures, Low-Level Programming, Database Systems, Algorithms, Cybersecurity, Artificial Intelligence, Machine Learning, Natural Language Processing, Data Science Techniques
- **Competitions**: Microsoft X PiE Datathon - Top 1 for best model, 27 teams total

Skills

- **Data Science**: Pandas, Scikit-learn, Pytorch, Matplotlib, Seaborn, NumPy, Regression, Neural Nets, NLP
- **Computer Science**: Data Structures, Algorithm Design, Low-Level Programming, Runtime Complexity, Databases, Cybersecurity Principles, RISC-V, Unit Testing, Jupyter, Docker, Git
- **Proficient Languages**: Python, Java, C/C++, SQL, Assembly, HTML, CSS

Work Experience

- **Student Researcher - Non-Orientable Manifold Editors (NOME)** **Fall 2020 - Spring 2021**
  - Student researcher with Professor Carlo Sequin to develop a new CAD tool, where I worked on generators (spherical, mobius strip, hyperboloid, general surfaces) and framework testing.
- **Undergraduate Student Instructor** **Fall 2020**
  - Worked under professors Fernando Perez and Anthony Joseph to facilitate an upper-division data science course to 1100+ students. Held remote discussion sections, labs, office hours, proctored & graded exams, and helped create course material (regression and modeling).
- **Data Science Intern at FreshLime** **Summer 2020**
  - Database deduplication with Tf-idf/levenshtein distance approaches, created a cohort retention interface generator for inclusion in the main product, presented chatbot latency analysis to the development head, and built new accounting software for VP of Business Development.
- **DataStory University Organization (Lead Consultant)** **Fall 2019 - Spring 2020**
  - Led student team on a client project working with environmental data, developed model and map-overlay visualization from raw datasets pulled from government sites.
- **IT Intern at AIDP Inc.** **Summer 2019**
  - Updated product websites (html/css) and implemented enterprise accounting software.

Projects

- **Neural Net Language Identification and Digit Classification**
  - Developed abstracted neural nets and optimized parameters (batch size, hidden layers, depth, etc.) to fit nonlinear functions, classify hand-drawn numbers, and determine language of input words.
- **NP-Hard Cell Towers - Minimum Weighted Connected Dominating Set**
  - Built a generator of graph edges for local minimal weight solutions on NP-Hard problem using random generation of MSTs on random dominating sets, before pruning on minimal weight MST.
  - Other approaches included probabilistic independent sets and Steiner Tree pruning.
- **Database Design**
  - Built B+ Tree Page Indexing, Relational Joins (Grace-Hash, Sort-Merge, etc.), Query Optimization, Concurrency Control (transaction locking), and Recovery Manager (ARIES).
- **Spam Email Detection Model**
  - Using a dataset of emails, built a logistic regression model in Python. The model was regularized and cross-validated, using cross-entropy loss. Achieved 95% test accuracy.
- **Python Library Matrix-Operation Speedup in C**
  - Developed Python matrix operation library in C using cache and SIMD techniques for speedups.
- **Playable Tile-Based Dungeon-crawler Game**
  - 2D array-based game with random seeded world generation, NPCs, saving, and win conditions.