D.Thong Nguyen (Pine Nguyen)

pinenguyen@berkeley.edu | (510) 345-7062 | https://www.linkedin.com/in/pine-ng

EDUCATION

University of California, Berkeley **Expected Graduation: May 2026**
Bachelor of Arts in Computer Science
Bachelor of Arts in Data Science

- **Activities and Societies**: ASUC OCTO - Software Engineer, Theta Xi Fraternity - Secretary
- **Relevant Coursework:** Data Structures and Algorithms, The Structure and Interpretation of Computer Programs, Linear Algebra, Discrete Mathematics and Probability Theory, Principles & Techniques of Data Science, Data Science for Economic Models, Foundations of Data Science, Deep Learning for Computer Vision, Concepts of Probability, Experimental Design

EXPERIENCES

Berkeleytime | Lead Backend Engineer *August 2024 - Current*

- Improving Berkeleytime.com, the most widely used indepdent platform at UC Berkeley, with 20,000+ monthly active users.
- Expanding the platform with crowd-sourced data and corrections. Adding functionality to rate courses, submit corrections for course-related data, and information about courses like past syllabi, whether attendance is required, etc.
- Specializing in application speed and mutation capabilities. Designed multiple aggregation layers, indexing, increases complexity in infrequent tasks to improve more frequent user executions.

Stackless Data | Software Engineering Intern *May 2024 - August 2024*

- Developed data pipelines to unlock valuable analytics for enterprise clients.
- Significantly reduced Snowflake compute cost (12% MoM) through more effective clustering and upgrades such as blocking simultaneous instances with shared dependencies, adding sensors for upstream changes, preventing the re-ingestion of fresh data.
- Responsible for transitioning the company's data orchestration stack to Dagster. Configured Dagster to replicate all existing functionalities and leveraged new capabilities afforded by the platform to improve efficiency.

PROJECTS

Pseudo-random generative world *August 2024*

- Developed a 2D tile-based world exploration engine. The project involved designing and implementing a system that generates pseudo-randomly created worlds consisting of interconnected rooms and hallways.
- Applied advanced data structures and algorithms to ensure efficient and scalable world generation.

Cross-Origin Data Injection Extension *August 2024*

- Created an open-sourced browser extension that integrates RateMyProfessors (RMP) statistics onto Berkeley's enrollment platform.
- Requested host permission privilege to bypass CORS, replicated interactions with RMP's internal GraphQL API through analysis of the network activity. Built an independent system to ingest, cache, and inject the data into the website.

Image Classification with CIFAR-10 Dataset *June 2024*

- Developed a Convolutional Neural Network (CNN) in PyTorch for CIFAR-10 image classification. The model incorporates convolutional, pooling, and fully connected layers.
- Achieved 96.2% test accuracy using only ~13,000 parameters. Optimization was performed using stochastic gradient descent and cross-entropy loss.

Linear Regression MLM for Property Tax Assessment *May 2024*

- Engineered features through data preprocessing and transformation techniques.
- Built a modular data processing pipeline to handle both training and test data scenarios, ensuring consistent feature engineering across datasets.
- Engineered features through data preprocessing and transformation techniques. Utilized data visualization (matplotlib/seaborn) to communicate model performance and identify potential biases in predictions.

SKILLS & INTERESTS

Languages/Tools: Java, Python, Javascript, Typescript, SQL, GraphQL, HTML/CSS, Scheme
Developer Tools: Git, Docker, GraphQL, DBT, MongoDB, Snowflake, Dagster, Jupyter Notebook
Libraries: Pytorch, Pandas, Numpy, Truth, Matplotlib
Interests: Audio, Graphic Design, Photography