

Thy Tran

U.S. Citizen | thyatran0113@gmail.com | +1 (669)-243-6747 | [linkedin.com/in/thyatran](https://www.linkedin.com/in/thyatran) | github.com/thyatran

Portfolio: thyt.netlify.app

EDUCATION

San Jose State University

San Jose, CA

Bachelor of Science in Data Science | GPA: 3.7/4

Expected Graduation: December 2026

- **Awards:** Dean's Scholar (2022), President's Scholar (2023)
- **Relevant Coursework:** Data Structures & Algorithms, Database Management Systems, Processing Big Data, NoSQL Database Systems, Discrete Math, Linear Algebra, Calculus I & II, Applied Probability & Statistics

Cornell University, online certificate in Machine Learning Foundations

May 2025 - August 2025

- **Relevant Coursework:** Exploratory Data Analysis, Machine Learning, Feature Engineering, Supervised & Ensemble Models, Model Evaluation & Deployment, Neural Networks, Convolutional Networks, Natural Language Processing

TECHNICAL SKILLS

- **Languages:** Python, Java, SQL, JavaScript, HTML, CSS
- **ML/AI:** Tensorflow, Keras, Scikit-learn, PyTorch, Pandas, NumPy
- **Visualization:** Tableau, Matplotlib, Seaborn
- **Tools/Cloud:** Git/Github, Jupyter, Google Colab, Excel
- **Database:** MySQL, MongoDB, Neo4j

EXPERIENCE

Automation Anywhere | Machine Learning Fellow (Break Through Tech AI Program)

Aug 2025 - Present

- Training **object detection models** (YOLO, Detectron2) to localize signatures and stamps in business documents
- Preprocessing datasets into **COCO/YOLO** formats to ensure balanced classes and accurate bounding boxes
- **Tuning hyperparameters** and comparing architectures to improve **precision, recall, and mAP** scores
- Building a **Streamlit** app for automated document data extraction

PROJECTS

Predicting Mental Health Model | Python, Scikit-learn, Git

2025 | [View Project](#)

- Built and compared **ML models** (Logistic Regression, Random Forest, MLPClassifier) to predict mental health needs
- Optimized hyperparameters with **GridSearchCV**, achieving **86% test accuracy** and improving model reliability
- Conducted **exploratory data analysis (EDA)** to identify key predictors of mental health outcomes in the workplace
- Presented findings through visualizations and insights to highlight industry trends and support awareness initiatives

Airbnb NYC Price Prediction | Python, Pandas, NumPy, Matplotlib, Seaborn, Scikit-Learn, Jupyter

2025 | [View Project](#)

- Built and optimized **regression models** to predict NYC Airbnb prices, explaining 60%+ of price variation
- Conducted exploratory data analysis (EDA) and **feature engineering** on 48,000+ NYC Airbnb listings
- Enhanced model through **cross-validation** and **hyperparameter tuning**, resulting in 15-20% performance boost

LEADERSHIP EXPERIENCE

Break Through Tech at Cornell Tech | AI/ML Fellow

May 2025 – Present

- Selected from 3000+ applicants for the Break Through Tech AI Program at Cornell Tech
- Participant in 12-month long program including Machine Learning coursework with Cornell faculty, experiential learning experiences, and mentorship from industry professionals
- Developed models in supervised/unsupervised learning, conducted data preprocessing, and data modeling

Girls Who Code SJSU | External Vice President

May 2025 - Present

- Secured industry partnerships, networking events, and career opportunities for student members
- Organized tabling and recruitment efforts, increasing membership sign-ups to 50+ students
- Coordinated with executive board to deliver engaging workshops and professional development opportunities