**Thu Nguyen**

\*US Permanent Resident\*

(971) 401 7226 | [tqn1472@gmail.com](mailto:tqn1472@gmail.com) | [github.com/tqn14](https://github.com/tqn14) | [linkedin.com/in/thuquynhnguyen](http://www.linkedin.com/in/thuquynhnguyen)

Personal website: [tqn14.github.io](http://people.tamu.edu/~tqn14/)

**EDUCATION**

**Texas A&M University** College Station, TX

B.S in Applied Mathematical Sciences, Statistics emphasis May 2021

Minor in Computer Science

Cumulative GPA: 3.62

**Related Coursework:**

Data Structures and Algorithms Linear Models Machine Learning

Principles of Data Sciences Statistical Computing Topological Data Analysis

**SKILLS and ACTIVITIES**

**Languages:** Fluent inEnglish and Vietnamese

**Technical Skills:** Proficient in Python (Pandas, Scikit-learn, Nltk, Framework: TensorFlow), R (ggplot2, dplyr, caret, RMySQL), C++; familiar with SQLite, Java

**Activities:** Pi Mu Epsilon, TAMU Data Science Club, Tethered Informatics and Data Analytics Lab

**EXPERIENCE**

**Texas A&M University** College Station, TX

*Research Assistant, Computational Materials Sciences Lab* Jan 2021 – Present

* Process 15 TB image data and apply Variational AutoEncoders for microstructural images extraction to better understand printing parameters and quality in 3D printing

*Undergraduate Researcher, Department of Petroleum Engineering* Aug 2020 – Jan 2021

"Data Driven Insights for Oil & Gas Drilling Operations"

* Applied Random Forest and Recurrent Neural Networks (in TensorFlow), with 99.48% accuracy rate and 85% precision rate in the first phase, to predict lithofacies of the wells and drilling dysfunctions on 2GB sensory downhole data
* Lead programming for a team of 3, assign programming tasks for each team member, oversee the process, and compile into a Jupyter Notebook

**San Jacinto College**  Pasadena, TX

*Research Assistant, Institutional Research & Data Science* Apr 2019 – Aug 2019

* Conducted phone surveys to registrar offices of 20 local high schools in Deer Park, South Houston, and Pasadena, presented the collected data in Excel spreadsheets in a clear manner for easy file access and data manipulation, and arranged the survey forms before sending them to other campuses

**LEADERSHIP**

*Peer Mentor, College of Science*  Aug 2020 – Dec 2020

* Co-instructed a first-year experience class, mentored, and met one on one biweekly with 6 freshmen

**PROJECTS**

Walmart Search Engine Optimization Oct 2020 – Present

(https://github.com/tqn14/Datathon2020-Walmart-search-engine-optimization-)

* Build a pipeline to crawl 20000 products information on walmart.com using Selenium in Python
* Create a GUI for a search engine built purely with Machine Learning tools, using TensorFlow Transform for text embeddings, SkLearn PCA for features reduction, and KMeans for clustering.