

Thu Nguyen

US Permanent Resident

(971) 401 7226 | tqn1472@gmail.com | github.com/tqn14 | linkedin.com/in/thuquynhnguyen

EDUCATION

Texas A&M University

B.S in Applied Mathematical Sciences, Statistics emphasis
Minor in Computer Science
Cumulative GPA: 3.59

College Station, TX
May 2021

Related Coursework:

Data Structures and Algorithms
Operational Research I, II

Mathematical Probability I, II
Principles of Data Sciences

Machine Learning
Statistical Computing

SKILLS and ACTIVITIES

Languages: Fluent in English and Vietnamese

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

Activities: TAMU Math Club, TAMU Data Science Club, Tethered Informatics and Data Analytics Lab

EXPERIENCE

Texas A&M University

Undergraduate Researcher, Department of Petroleum Engineering
"Data Driven Insights for Oil & Gas Drilling Operations"

College Station, TX
August 2020 – Present

- Filter high frequency well field data from 2000Hz to 10Hz using a Fast Fourier Transform (fft) in R and down sample the data with Interpolate in Python Pandas
- Apply Random Forest and Artificial Neural Networks to predict geological features of the wells and drilling dysfunctions, with 99.48% accuracy rate and 85% precision rate in the first phase
- Lead programming for a team of 3, assign programming tasks for each team member, oversee the process, and compile a report every other week

Peer Mentor, College of Science

August 2020 – Present

- Meet one on one with 6 assigned freshmen to facilitate their personal and academic transitions to Texas A&M University throughout the semester
- Plan course components with the instructor to encourage the students' participation and direct class activities in a breakout room

San Jacinto College

Research Assistant, Institutional Research & Data Science

Pasadena, TX
April 2019 – August 2019

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

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
- Apply TensorFlow Transform for text embeddings, sklearn PCA for features reduction, and KMeans for clustering. Create a GUI for queries on a web page that hosts this search engine