

Khuyen Tran

[LinkedIn](#) | [Github](#) | [Website](#) | [Articles](#) | Edwardsville, Illinois | (618)514-1815 | khuyentran1476@gmail.com

EDUCATION

Southern Illinois University Edwardsville (SIUE)
Bachelor of Science in Applied Math with Honors
Mathematical Tools for Data Science in CIMAT

Expected Grad: May 2021
Cumulative GPA: 4.0 / 4.0
Spring 2020

- **Related courses:** Probability and Statistics, Linear Algebra, Design of Algorithms, Discrete Mathematics, Numerical Optimization and Machine Learning, Data Visualization and Computational Geometry, PostgreSQL

SKILLS & EXPERTISE

- **Programming languages:** Python, R, SQL, C++
- **Libraries & Frameworks:** Keras, Pandas, Tensorflow, Numpy, Scikit-learn, Matplotlib, NLTK, Integer Programming
- **Tools:** Talend ETL, AWS RDS, Tableau
- Strong knowledge in Machine Learning, Data Science, Data Analysis, Data Visualization and Natural Language Processing.

PROJECTS

[Scheduling App with Machine Learning and Integer Programming](#)

Jan 2020-Present

- Cooperate with Dr. [Andrew Trapp](#) from Worcester Polytechnic Institute (WPI) to develop a scheduling app
- Utilize integer optimization to find the most effective schedule and maximize productivity.
- Tools: integer programming with Pulp-optimization in Python, Gurobi and natural language processing for task classification

[Predict Gender and Language Variety in Twitter](#)

March 2020

- Identify author's gender and language variety in Twitter.
- Experiment with 3 different ML models and achieve f1-score of up to .81 for gender prediction and .84 for language variety.
- Tools: NLTK (natural language toolkit) for preprocessing raw texts, scikit-learn (machine learning in Python), BERT (natural language processing), and neural network for model training and prediction.

[Scraping, Text Processing, and Analysis Ghibli Movie Database](#)

Jan 2020

- An end-to-end exploratory data analysis: extracting raw data on Ghibli Movie Database and analyzing the data.
- Tools: Beautiful Soup (Python library for web scraping), NLTK for text processing, Pandas and Matplotlib (Python library for visualizing and analyzing the results)

EXPERIENCE

Data Science Intern

March 2020-Present
St. Louis, MO

Arch Talents

- Utilize ETL to pull data from its source, prepare queries to transform data into the AWS data warehouse.
- Process data and create predictive model with Python.
- Deploy the model and visualize with Tableau.

Research Assistant

Jan 2019-Dec 2019
Edwardsville, IL

SIUE Department of Mathematics and Statistics

- Formulated metaheuristic with Dr. Chew, a method to provide optimal solution with incomplete information within acceptable time. Especially useful for neural network training
- Achieved 10-6 accuracy when tested on multidimensional functions

Data Science Technical Writer

Jan 2020-Present

[Towards Data Science](#)

- Authorized more than 30 articles on topics of Natural Language Processing, web scraping, data science tools, and mathematical programming with more than 110k views a month
- Simplified complex mathematical and programming concepts with insights and interactive visualizations.

Math and Business Tutor

Aug 2018- Dec 2019
Edwardsville, IL

SIUE Tutoring Resource Center

- Guided more than 10 students to get from C to A grade in statistics and mathematics courses within 2 months.
- Effectively communicate to those who have little understanding of the subjects with confidence and clarity.

Calculus Enrichment Session Leader (ES Leader)

Aug 2018- Dec 2018
Edwardsville, IL

SIUE Department of Mathematics and Statistics

- Led a class of 60+ students from different learning backgrounds to master calculus knowledge.
- Devised strategic plans with other leaders and professors to deliver complex concepts in comprehensible ways