

Yu Fang

Pittsburgh, PA | yufang@andrew.cmu.edu | +1 (412) 969 9213 | <https://www.linkedin.com/in/lucia-fangyu/>

EDUCATION

Carnegie Mellon University

Bachelor of Science | Information Systems

CMU Dietrich College Dean's List with High Honors 2021-Present

May 2025

PUBLICATIONS

Fang, Yu, Huang, Shihong, & Ogan, Amy (Submitted, International Learning Analytics & Knowledge Conference). A Cross-Cultural Confusion Model for Detecting and Evaluating Students' Confusion In a Large Classroom.

Fang, Yu (Lucia), Patidar, Prasoon, Zimmerman, John, Agarwal, Yuvraj, & Ogan, Amy (2024). Engaging or Zoning Out in Class: Automated Engagement assessment with Unsupervised Clustering.

Gersey, Julia, Gui, Angela, **Fang, Yu (Lucia)**, Ngoon, Tricia, Patidar, Prasoon, Zimmerman, John, Agarwal, Yuvraj, & Ogan, Amy (2023). What's happening in the classroom? Automated recognition of classroom activity for scalable multimodal learning analytics.

RESEARCH EXPERIENCE

Senior Honor Thesis, Carnegie Mellon University

A Study of Student Facial Expression for Revealing Difficult Topics In Class (Manuscript, In Preparation)

May 2024 - Present

- Developed a cross-cultural Convolutional Neural Network (CNN) confusion model.
- Implemented the confusion model within DeepFace and analyzed 15 classroom videos (~600K facial images).
- Used SHapley Additive exPlanations (SHAP) to validate racial differences in literature.
- Implemented audio-based Retrieval Augmented Generation (RAG) and developed user interface via Streamlit for 15 sessions.

Small Undergraduate Research Grants, Carnegie Mellon University

Engaging or Zoning Out in Class: Automated Engagement assessment with Unsupervised Clustering (Poster)

January 2024 - May 2024

- Preprocessed and clustered eye tracking features with UMAP and k-means into attention motifs across 25+ STEM sessions.

Undergraduate Research Assistant, Carnegie Mellon University

What's happening in the classroom? Automated recognition of classroom activity for scalable multimodal learning analytics (Poster)

November 2022 - December 2023

- Trained and compared Logistic Regression and Random Forest models for 15 classroom activities across 25+ STEM sessions.
 - Resolved data imbalances using techniques like ADASYN and SMOTE.
-

WORK EXPERIENCE

Technical Intern, Siemens Innometrics

- Developed 20+ Qlik dashboards for business stakeholders using SQL in Snowflake.

June 2023 - Present

- Preprocessed and generated insights to facilitate stakeholders' decision making using Python.

- Implemented Retrieval Augmented Generation (RAG) to help users interpret document content in Snowflake Streamlit.

PROJECT EXPERIENCE

Music Recommendation for Mental Health, Carnegie Mellon University

- Performed exploratory data analysis, data preprocessing, and feature selection on music preferences for mental health conditions.

January 2023 - May 2024

- Applied machine learning to predict mental health conditions to highlight the potential of personalized music therapy as a supplement to traditional treatments.

Customer Spending Behavior Analysis, Carnegie Mellon University*January 2023 - May 2024*

- Conducted time series analysis and PCA to identify key predictors of spending.
- Analyzed education and marital status with respect to spending behavior to offer insights for targeted advertising.

Web Development for Crime Tracker, Carnegie Mellon University*January 2023 - May 2024*

- Designed and developed a crime tracking app using Ruby on Rails with MVC architecture and React.
- Wrote unit tests for MVC components and conducted testing with fabricated datasets.

Social Media User Relationship Modeling, Carnegie Mellon University*November 2023 - December 2023*

- Implemented and fine-tuned a PostgreSQL database.
- Employed advanced SQL and Python scripting for robust data management and application functionality.

Mobile App for Therapy, Carnegie Mellon University*January 2023 - May 2023*

- Developed a real-time multimodal and multi-language emotion recognition app with DeepFace, SpeechRecognition, revChatGPT.
- Deployed the application on Streamlit Cloud.

TEACHING EXPERIENCE

Database Design and Development (67-262), Carnegie Mellon University

August 2024 - Present

Digital Accessibility(67-220), Carnegie Mellon University

August 2023 - December 2024

Principles of Computing (15110), Carnegie Mellon University

July 2023 - August 2023

SKILLS

Languages: Python, Java, Javascript, React.js, C++, PostgreSQL, MongoDB, R, Spark, Ruby, Swift**Packages/Applications:** Numpy, Pandas, Seaborn, Scipy, Streamlit, Scikit-learn, Tensorflow, Snowflake, Arduino IDE, NodeBB**Tools:** Figma, Qlik Sense**Design skill:** Wireframing, User Flow, Prototyping, Data Visualization**Research:** Interviews, Quantitative and Qualitative data analysis**Communications:** English (Native / Bilingual), Mandarin (Native / Bilingual), Spanish (Advanced)