Jacob Tae H. Emmerson

Pittsburgh, PA • emmerson.jacob@gmail.com • 563 726 9927 • jacobemmerson.github.io/

Research Interests: A.I. for social good (AI4SG), representation learning with multiple modalities, common sense and reasoning, concept learning by induction

EDUCATION

University of Pittsburgh, Pittsburgh Campus — Aug. 2021 - May 2024
School of Computing and Information

- B.Sc. Computer Science; *summa cum laude*
- Minor in Applied Statistics

Graduate Coursework: Foundations of A.I., Advanced Topics in A.I.

RESEARCH EXPERIENCE

Research Assistant, PI: Ryan Shi — Jul. 2024 - Present
University of Pittsburgh, **AI 4 Social Good**

- Designed a retrieval-augmented generation (RAG) pipeline using Google Search and Semantic Scholar APIs to enhance societal awareness and understanding.
- Developed a problem scoping agent with the retrieval pipeline and a large-language model (GPT-4o) to propose literature-backed applications of AI on complex societal challenges.
- Project proposals composed by the full framework were determined to be a higher quality than one-shot generations according to blind reviews and AI-agent evaluations; *paper in progress*.

Research Assistant, PI: Adriana Kovashka — May 2024 - Aug. 2024
University of Pittsburgh, **Deep Learning and Computer Vision**

- Validated the existence of measurable differences in perception across cultures, inspired by observations in cognitive psychology, using CrossModal-3600 and Multi30k.
- Investigated differences in the attention distributions of cross-lingual vision-encoders in multimodal models.
- Evaluated the cross-cultural reasoning of LLMs and VLMs in low-resource domains using few-shot and in-context learning through caption rewriting (BLEU, ROUGE)

INDUSTRY EXPERIENCE

Machine Learning Developer — Aug. 2022 - Present
Signature Diagnostics

- Proposed transformations inspired by log-ratio analysis to learn genetic features (RNA-Seq, Methylation) resistant to batch-effects.
- Created an ensembling framework to improve prenatal disease classification accuracy by quantifying the uncertainty of individual classifiers.

SKILLS

Programming Languages: Python, R, C/C++, Java, MATLAB
Preferred Libraries: PyTorch, PyG, Pyro, HuggingFace, NLTK, Pandas, NumPy, Seaborn
Technical Skills: Bayesian statistics, probabilistic programming, multivariate and non-parametric statistics, large-language models (LLMs), vision-language models (VLMs), natural language processing (NLP), computer vision (CV), multimodal learning, data science, graphical modeling, sequential modeling
Soft Skills: Collaboration and teamwork, scientific writing, project management, team leadership and advocacy

ORGANIZATIONS

Rainbow Alliance, Board Member — Oct. 2021 - Apr. 2022
Student Government Board, Judicial Committee Member — Apr. 2022 - Dec. 2022

PROJECTS

Genetic Algorithm for Equitable Neighborhood Service Mar. 2024
Python
- Encodes localized information (grocery stores, retail stores, distances) obtained from Google Map's API using a multi-layered perceptron.
- Optimizes a weighted transit network using evolutionary algorithms and the encoded states at each bus stop.

Quantum Hadamard Edge Detector Apr. 2023
Qiskit
- An edge detection algorithm for image analysis utilizing quantum gates implemented for comparison against classical alternatives; developed and tested remotely using IBM's Quantum Computers.
- https://github.com/jacobemmerson/QHED

Textual Entailment Model for Question Answering Dec. 2023
Python
- An RTE-based model for answering multiple-choice questions about a given set of text; trained and evaluated on the publicly available MC500 dataset.
- https://github.com/jacobemmerson/CS1671/tree/main/MCTest