Jacob T. Emmerson

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

EDUCATION

University of Pittsburgh, Pittsburgh Campus — Aug. 2021 - May 2024
School of Computing and Information
- B.Sc. Computer Science; *summa cum laude*
- Statistics Minor

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

INDUSTRY EXPERIENCE

Machine Learning Developer — Aug. 2022 - Present
Signature Diagnostics
- Created a Bayesian ensembling framework that increased accuracy by $\sim 5\%$ on non-invasive prenatal disease classification.
- Recommended several innovative ideas for understanding the nuanced differences between sequencing techniques.
- Provided intuitive explanations and recommendations of statistical methods for interdisciplinary research between biologists and engineers.
- Developed methods inspired by log-ratio analysis for creating features resistant to batch-effects.

ACADEMIC EXPERIENCE

Research Assistant, PI: Ryan Shi — Jul. 2024 - Present
University of Pittsburgh, **AI 4 Social Good**
- Designed a retrieval-augmented generation (RAG) framework using Semantic Scholar's API to enhance understanding of domain-specific challenges for public sector organizations.
- Applied GPT-4o to identify innovative AI-applications of previous research tailored to public sector needs.
- Conducted in-context learning experiments to improve the adaptability of AI models in real-world, domain-specific scenarios.

Research Assistant, PI: Adriana Kovashka — May 2024 - Aug. 2024
University of Pittsburgh, **Deep Learning and Computer Vision**
- Investigated biases in attention mechanisms of cross-lingual vision-encoders in vision-language models (VLMs).
- Evaluated the cross-cultural reasoning of LLMs and VLMs in low-resource domains using in-context learning and retrieval augmented generation.
- Proposed a fine-tuning approach to efficiently improve the cross-cultural awareness of language models.

SKILLS & INTERESTS

Research Interests: A.I. for social good (AI4SG), grounded representation learning in multimodal models, reasoning in generative models, neuro-symbolic approaches to AGI
Programming Languages: Python, R, C/C++, Java, MATLAB
Preferred Libraries: PyTorch, PyG, PyBBN, HuggingFace, NLTK, Pandas, NumPy, Seaborn
Technical Skills: large-language models (LLMs), vision-language models (VLMs), natural language processing (NLP), computer vision (CV), mulitmodal learning, data science, distributed training, graphical modeling, sequential modeling, multivariate and non-parametric statistics
Soft Skills: Collaboration and teamwork, interdisciplinary research, scientific writing, project management, team leadership and advocacy

ORGANIZATIONS

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