University of Pittsburgh | School of Computing and Information

Proposal Defense

Doctor of Philosophy in Computer Science

"**Understanding, Evaluating, and Generating Advertisement**" by **Aysan Aghazadeh**

Date: Tuesday, June 23, 2025
Time: 3:00 – 5:00 p.m.
Place: 6143 Room, Sennott Square, 3810 Forbes Ave, Pittsburgh, PA 15213

Committee:

- **Adriana Kovashka**, Associate Professor, Department of Computer Science, University of Pittsburgh
- **Diane Litman**, Professor, Department of Computer Science, University of Pittsburgh
- **Xiang Lorraine Li**, Assistant Professor, Department of Computer Science, University of Pittsburgh
- **Pinar Yanardag**, Assistant Professor, Department of Computer Science, Virginia Tech

Abstract:
Advertising media plays a significant role in shaping the daily decisions and behaviors of people around the world. However, advertisements are rarely effective if they merely depict objects in a straightforward manner. To enhance persuasiveness, designers often employ creative strategies and persuasive techniques, such as atypical object presentations or the integration of cultural elements. With recent advances in generative AI, the use of these models in media generation and interpretation has grown. However, the creativity and persuasion demand of advertising content make unique challenges for both the generation, and interpretation of such media.

In this proposal, I first investigate current shortcomings in the generation and interpretation of advertisement media. I demonstrate that while Multimodal Large Language Models (MLLMs) perform well on standard tasks, they fall short in interpreting and evaluating content that relies on creativity and persuasion. Similarly, state-of-the-art Text-to-Image (T2I) generation models, despite producing high-quality and realistic images, often fail to create content that is persuasive or memorable.

To address these limitations, I propose a reinforcement learning with human feedback (RLHF) approach to train the projection layer in MLLMs, aiming to produce more accurate visual embeddings and enhance reasoning capabilities, particularly for atypical or unusual images. Furthermore, I propose leveraging style images and a symbolism knowledge base to guide the generation of visually atypical images that evoke specific physical sensations. This involves selecting symbolic objects that best represent the targeted sensation and, using a large language model, identifying the most effective atypical configurations. A learned representation of object relationships from example images with similar atypicality will be used as an additional conditioning signal.

Additionally, I propose the creation of a human-annotated dataset evaluating memorability and persuasiveness, which will serve as an RLHF reward signal to fine-tune T2I models for generating impactful advertisement content. Finally, I extend this work to the generation of advertisement videos that are both persuasive and creatively engaging.