Ph.D. in Information Technology
Thesis Defense

May 18th, 2023
at 15:30
Aula PT1

Vincenzo SCOTTI – XXXV Cycle

GENERATIVE EMPATHETIC DATA-DRIVEN CONVERSATIONAL AGENTS FOR MENTAL HEALTHCARE

Supervisor: Prof.ssa **Licia Sbattella**

Abstract:

Building intelligent artificial agents, able to hold conversations with humans, has been a long-running goal of artificial intelligence.
In this sense, affective computing introduced the idea that, for a machine to be perceived as intelligent, it should be able to understand emotions.
This idea has been expanded to include the broader concept of empathy, which is the ability to understand and share what other individuals are experiencing. Empathy is a complex phenomenon, which requires the use of both cognitive and emotional intelligence abilities, and affects both segmental (i.e., words and sentence structures) and suprasegmental (i.e., how the sentence is spoken) levels of human conversations.

In this research, we focus on developing empathetic conversational agents (i.e., conversational agents capable of simulating empathy) for mental healthcare.
We developed multiple agents for applications that require an empathetic approach, from plain open-domain dialogues to more complex interactions, like therapy sessions.
In particular, the latter represents the final objective of our work.

Embodying conversational agents, through voice or avatars, is known to make them more relatable, thus improving their perception as more human and intelligent. Therefore, apart from a simple text-based interface, we also provided modules for spoken input and output to complete our agents.
In particular, trying to simulate the suprasegmental effect of empathy on speech, we augmented the agent's vocal synthesis with a module that adapts its speaking style depending on the conversation status.

Given the latest results in Natural Language Processing brought by deep learning-based solutions, we decided to design our agents focusing on such data-driven solutions.
We developed multiple empathetic dialogue agents using different learning paradigms, like reinforcement and curriculum learning, and different models, like latent hierarchical and prompt-based models.
Some of such agents are capable of conditioned text generation on aspects like dialogue acts and emotions, as well as recognition of these same aspects, helping explain and understand the output they produce.

We evaluated our agents using both human and automatic approaches, obtaining promising results. Given the observed outcomes and the latest developments in NLP yield by (deep learning-based) large language models, we expect that scaling up the underlying dialogue model complexity would improve significantly the agents' capabilities.
Finally, we deployed a demo agent for therapy, on an instant messaging application.
This last deployment step allowed us to investigate whether it could give users easy access to these conversational agents without the need for complex web application pipelines.

PhD Committee
Prof. **Mark James Carman,** Politecnico di Milano
Prof.ssa **Barbara Di Eugenio**, University of Ilinois Chicago (UIC)
Prof. **Najim Dehak**, Johns Hopkins Whiting School of Engineering