Reason, Advise, and even Mind-Reading: Multiple Applications of Large Language Models

Daniel Hajialigol, Hanwen Liu, Meng Lu, Zhenyu Bi, Xuan Wang

Department of Computer Science, Virginia Tech

Fundamental NLP

Weakly-supervised Information Extraction:

- **Background and Prior Work:** State-of-the-art information extraction techniques for tasks such as named entity recognition, entity typing, text classification, relation extraction, etc. typically involve massive amounts of labeled data; however, this could be intractible to do in certain domains such as medicine and cyber. To overcome these issues, many works attempt to operate in the weakly supervised scenario, where few-to-no ground truth labels are provided.
- **Method and Results:** We have worked on an extremely-weakly supervised text classification algorithm, XAI-Class, that explicitly models salient, or important, tokens in the input via a multi-task model. XAI-Class performs as or very close to state-of-the-art on experimented datasets, indicating the utility of explicitly modeling important tokens in the input.
- **Future Work:** Our future work would be to extend this concept to multiple modalities. Additionally, it is a goal to extend this framework to different information extraction tasks (entity recognition/typing, relation extraction, event extraction, etc.).

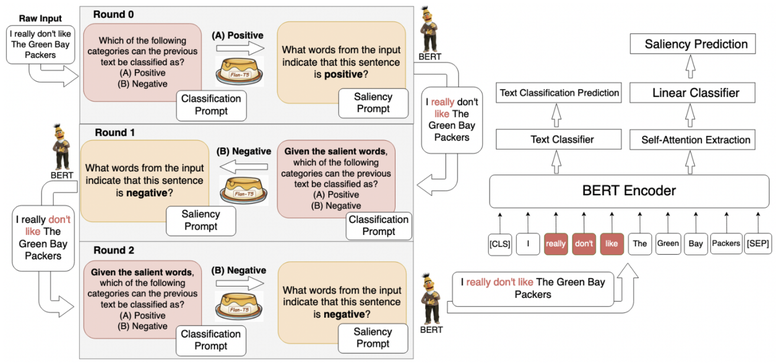

Figure 1. Proposed Architecture for Weakly-Supervised IE

Trustworthy Multi-Hop Question Answering:

- **Backgound:** Multi-hop question answering (MHQA) requires a model to retrieve and integrate information from multiple passages to answer a complex question. Recent systems leverage the power of large language models and integrate evidence retrieval with reasoning prompts (e.g., chain-of-thought reasoning) for the MHQA task.
- **Challenges:** The complexities in the question types (bridge v.s. comparison questions), as well as the reasoning types (sequential v.s. parallel reasonings), require more novel and fine-grained prompting methods to enhance the performance of MHQA under the zero-shot setting.
- **Method:** Inspired by the Tree-of-Thought (ToT) prompting method, we propose a ToT-based MHQA method that allows the model to generate different reasoning paths from the same question, thus effectively avoiding reasoning dead-ends.
- **Future Work:** We are currently working on refining tree-branching procedures and evaluating criteria. A wider goal would be solidify uncertainty quantification and provide better datasets for evaluation.

Fundamental NLP

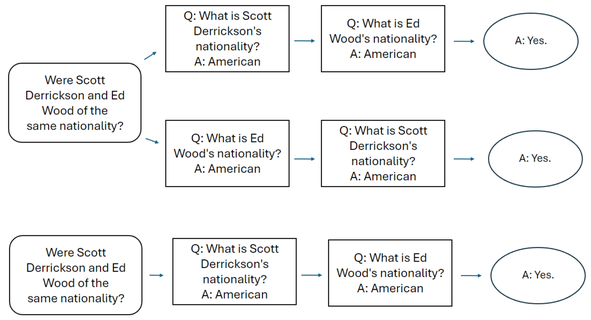

Figure 2. Tree-of-Thought (Top) and Chain-of-Thought (Down) reasoning paths

Biomedical NLP

- **Challenges:** At present, two primary obstacles prevent LLMs from accomplishing tasks related to medicine:
  1. The quantity and specificity of training data in the medical field are limited compared to the general web data used for training LLMs.
  2. Achieving high performance in this field necessitates extensive domain knowledge and advanced reasoning skills.
- **Prior Work:** To bridge this gap, there is a growing trend of methods aiming to equip LLMs with improved proficiency in medical knowledge through instruction tuning. These methods rely on either external knowledge bases or self-prompted data to create instruction datasets. However, acquiring high-quality instruction tuning data in the medical domain is costly, subject to privacy concerns, and thus not scalable.
- **Methods:** We have developed a multi-agent framework in the clinical domain, aimed at boosting the LLM's capacity for clinical text classification tasks through interactive scenarios.

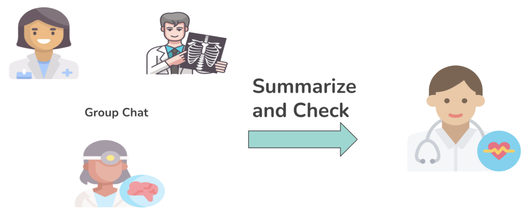

Figure 3. Overview of the Framework. Four LLM agents each have one different role.

- **Preliminary Results:** With our preliminary design of multi-agents, the total discordance of Clinic Records Text Classification Predictions decreases from 48.6% to 39.6% with GPT3.5 as the backbone model.
- **Future Work:** For future steps, we plan to design different schemas on both the agents' side and the LLMs' side. We will incorporate more Customized LLMs such as ClinicGPT, and teachable models developed by us. Simultaneously, we will introduce more external resources as domain experts for agent design.

Brain-to-Text

- **Background and Challenges:** The field of brain-computer interfaces (BCI) is advancing rapidly with the aim of facilitating communication and interaction for disabled individuals and enhancing virtual reality experiences.
  EEG-based research faces challenges due to lower resolution and limited dataset sizes, hindering model accuracy and generalizability.
- **Model Overview:** Our proposed foundational model for diverse brain EEG data representation aims to overcome the challenges associated with EEG signal processing and model generalizability, addressing the limited size of EEG datasets and the variations in data formats across different datasets.
  The model operates within a unified framework, integrating various EEG data formats and leveraging pre-training techniques to construct comprehensive representations that transcend dataset-specific idiosyncrasies.

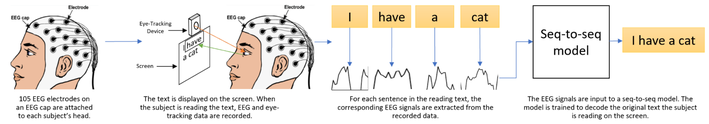

Figure 4. Overview of the Framework.

- **Specific Tasks:** We propose four tasks within this model.
  Task 1 involves establishing a unified preprocessing framework and dataset development by merging EEG datasets based on electrode placement distribution.
  Task 2 focuses on encoding spatial-temporal information into the model to enhance feature representation capability.
  Task 3 employs pre-training techniques tailored to EEG data characteristics to enhance encoding capabilities. This includes temporal and spatial pre-training to capture meaningful temporal and spatial representations of brain activity.
  Task 4 explores downstream applications such as text decoding, image decoding, emotion classification, and zero-shot image-to-text translation bridged by EEG, showcasing the versatility and utility of our foundational model across a wide range of brain-computer interface applications.

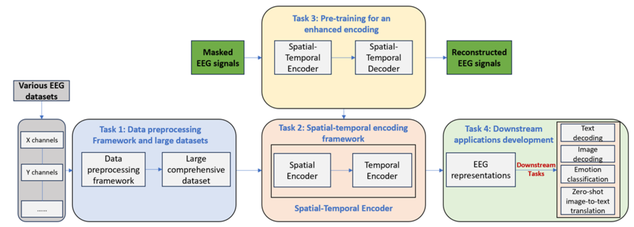

Figure 5. Specific Tasks in the Framework

Acknowledgement

Our work is sponsored by the Commonwealth Cyber Initiative, Children's National Hospital, Fralin Biomedical Research Institute (Virginia Tech), Sanghani Center for AI and Data Analytics (Virginia Tech), Virginia Tech Innovation Campus, and a generous gift from the Amazon + Virginia Tech Center for Efficient and Robust Machine Learning.

Visit our lab website for more details!