ABSTRACT

Human communication has been fundamental to societal development since the dawn of civilization, evolving from primitive gestures to complex languages. Modern technological advancements have revolutionized communication, enabling instant global interactions and sophisticated AI-driven intent detection systems. Despite these advancements, issues in human-machine communication persist, particularly in understanding and processing vast amounts of data. Addressing fundamental natural language understanding issues related to reasoning, comprehension, and learning is crucial for enhancing the effectiveness of these interactions.

To improve communication between humans and machines, integrating intent detection and key information generation is vital. In the modern digital world, vast external knowledge bases necessitate the efficient identification and extraction of relevant information to ensure precise and contextually appropriate responses. By summarizing and focusing on key data points, machines can better understand user intent and provide more accurate answers without being overwhelmed by extraneous details. This targeted approach not only enhances the efficiency of information retrieval but also improves the quality of machine-generated responses, leading to more natural and productive interactions. As NLP research progresses, refining architectures to better emulate human conversation remains a key focus for advancing machine comprehension and response generation.

Intent Detection has been a major field of study for the past few years, focusing on its various characteristics. Identifying user intents from text or voice is a crucial aspect of conversational dialogue systems. To interact effectively with users and provide appropriate responses, these systems must parse and interpret language expressions, particularly the users' intentions and identify the key contents. How-

ever, detecting intent dynamics from text or voice can be challenging and it is even harder for various domain-specific applications especially in India, where multiple languages are spoken concurrently, and multilingual syntaxes appear intertwined in a single conversation.

In this work, we aim to tackle the nuanced task of identifying intents, extracting intent with corresponding entities and presenting vital information across various domains. Beginning with the enhancement of identifying multiple intents from task-oriented dialog systems, we check it's effectiveness on various domain-specific applications. Then the focus shifts towards generating key contents which include persona-based summarization and highlight generation.

Voice-command based conversational agents typically rely on a predefined set of skills or intents to perform user tasks. As new intents emerge over time, these systems need to be retrained to handle them. This process involves two key challenges: identifying new intents dynamically and annotating data for these intents to retrain the classifier efficiently. These challenges are amplified when many new intents arise simultaneously and manual annotation resources are limited. To address this, we propose Multiple Novel Intent Detection (MNID), a cluster-based framework designed to detect and annotate multiple new intents with a constrained budget for human annotation.

In our exploration of the domain-specific application of intent detection, we focus on multilingual Indian healthcare scenarios. We perform extensive experiments on a set of models in various realistic settings to assess the current situation with its effectiveness in India and propose two datasets in seven different Indian languages with healthcare intents and corresponding entities.

Next, we investigate the critical issue of key content generation in the form of persona-based summarization and highlight generation. Persona-based summariza-

tion further refines the extraction process, catering to the diverse information needs of different stakeholders within specific domains like healthcare (doctors, patients, etc.). Highlight generation emerges as another crucial aspect of this endeavor, offering a streamlined method for generating the most impactful and inquisitive insight from voluminous texts. By automatically highlighting significant points, this process not only aids in comprehension but also fosters curiosity among readers, compelling them to delve deeper into the material. Through meticulous evaluation and experimentation, the study demonstrates the effectiveness of fine-tuning and preference-based optimization approaches in generating key contents. By leveraging AI-based critiquing, the approaches ensure the accuracy and scalability of generated key information, addressing the inherent variability associated with human-generated content.

Overall, this holistic approach to intent detection and key content generation underscores the importance of leveraging cutting-edge techniques to address the evolving challenges of information retrieval and presentation in domain-specific applications.

Keywords: intent detection, novel intents, health-care intents, entity extraction, healthcare entity, Indian health-care, persona specific summary content generation, key information, information extraction, highlight generation, abstract highlight, voice assistant, search, few-shot, zero-shot, silver annotation, LLM, LLM data generation, GPT-4, Gemini, small sized LLMs, LLM as evaluation, GPT-4 critique, Gemini critique, budget annotation.