User Intent Recognition using Natural Language Processing and Large Language Model

Prashant Mittal, Shyam Prajapati, Keshav Gulati
Indian Institute of Technology Ropar
Department of Computer Science Engineering
Rupnagar, Punjab, India
{2020csb1113, 2020csb1110, 2020csb1095}@iitrpr.ac.in

ABSTRACT

In the evolving landscape of customer service, shaped by the transformative impact of the Internet, the imperative to decode user intent in the realm of customer executive tasks has become increasingly paramount. This abstract outlines a solution that addresses the multifaceted challenge of discerning user intent, where queries exhibit diversity and implicit expressions. The proposed framework revolves around the development of a precise AI model, harnessing advanced Natural Language Processing (NLP) techniques and trained on diverse intent datasets.

Our model introduces an interactive approach to streamline user intent detection. Through targeted questions, the model iteratively refines its understanding based on user responses, progressively narrowing down the set of potential user intents. Unlike traditional approaches that often rely on static models, our method embraces an adaptive and interactive framework, aligning with the dynamic nature of customer queries using a Large Language Model(LLM). Responses can be generated for several intents at a time using LLMs. The proposed solution ensures a more responsive and effective customer support experience across diverse contexts.

Transitioning from the abstract to the term paper's objectives, our aim is to design and develop a question-answer system that accurately determines user intent in the customer executive task. Intent classification, facilitated by machine learning algorithms, categorizes user queries into different intent classes. The system, trained on a labeled dataset, predicts the intent of new user queries.

Upon intent determination, the system generates answers using a combination of rule-based and machine learning-based approaches. Rule-based methods handle common queries, while machine learning-based approaches address complex and context-dependent ones.

Testing on a diverse set of user queries assesses its effectiveness in accurately determining user intent and providing relevant answers. This term paper contributes to question-answering systems by proposing an efficient and accurate system for determining user intent in customer executive tasks, potentially enhancing customer support services through automated understanding and prompt, accurate responses.

I. Introduction

Transforming User Interaction using the *Dynamic Power* of Language Models. Gone are the days of rigid, pre-defined queries for specific intents. Thanks to Language Models (LLMs), chatbot interactions have evolved into dynamic, creative exchanges. LLMs enable the generation of spontaneous and diverse queries, not just for one intent but for multiple intents simultaneously. This shift from a scripted to a dynamic conversation transforms the user experience, making the chatbot feel less like a static machine and more like an intelligent companion. The magic lies in the adaptability and creativity of LLMs, turning each interaction into a linguistic symphony that captivates users and keeps them engaged. Welcome to an era where every conversation is an exploration, and the chatbot is your ever-evolving conversational partner in the world of language.

In the contemporary landscape of customer support services, the role of efficient and personalized interactions between users and customer executives cannot be overstated. One of the critical elements in achieving this lies in the accurate discernment and understanding of user intent. As the demand for seamless communication channels continues to grow, there is an imperative need for robust and dynamic User Intent Detection (UID) systems.

This paper introduces a comprehensive approach to User Intent Detection, specifically tailored to address the intricate tasks handled by customer executives. At the heart of our proposed solution is a sophisticated AI model, meticulously crafted and trained on diverse intent datasets, incorporating advanced Natural Language Processing (NLP) techniques. The central objective is to streamline the user-query resolution process by introducing a novel feedback loop, iteratively narrowing down potential intents and improving the precision of customer executive responses.

Crafting a Precise AI Model: Our methodology begins with the initialization of a set containing all conceivable user intents. This set acts as a comprehensive starting point for the model. The model then employs NLP text matching, dynamically adjusting the set based on a probability threshold derived from initial user input. This initial filtering not only refines the set of potential intents but also establishes a foundation for an optimized Response Generation phase.

Dynamic NLP Text Matching: The NLP text matching phase is not static; it dynamically adapts the set of potential intents throughout the user interaction. By constantly reassessing and filtering intents based on matching probabilities, the system ensures that only the most relevant and likely intents persist. This adaptive filtering mechanism enhances the precision of subsequent stages, laying the groundwork for more accurate user intent recognition.

Response Generation with Language Model (LLM): In response to the dynamically filtered intent set, our system utilizes a Language Model (LLM) to generate contextually relevant and coherent responses. This integration ensures that customer executives are equipped with meaningful suggestions, facilitating a smoother and more effective interaction with users. The LLM adapts its responses based on the evolving understanding of user intent, contributing to a more responsive and user-centric support experience.

User Query Processing and Feedback Loop: The system's evolution doesn't end with the initial response. The User Query Processing phase continuously processes user feedback, refining the intent set further. To enhance user engagement and intent confirmation, an Intent Confirmation Loop is introduced. This loop generates a strategic set of questions designed to ascertain and confirm user intent. The responses to these queries not only guide the system in confirming the user's intent but also contribute to the ongoing learning process, enabling the model to adapt to evolving user needs.

Final Intent Confirmation: The culmination of our approach involves a Final Intent Confirmation step. If only a few intents remain after the confirmation loop, they are considered the final user intent. In cases where a substantial number of intents persist, the system intelligently seeks additional clarification from the user. This iterative process ensures a comprehensive understanding of user requirements, paving the way for more accurate and personalized support.

In summary, this paper puts forth an innovative and adaptive framework for User Intent Detection, tailor-made for the intricate tasks handled by customer executives. By integrating advanced NLP techniques, a dynamic feedback loop, and a responsive Language Model, our approach aims to elevate the efficiency and accuracy of customer support services, contributing to a more seamless and satisfying user experience in the realm of customer interaction and assistance.

II. System Model

1. Data Collection and Training:

Dataset: To facilitate the development of a robust intent recognition model, we began by assembling a diverse and comprehensive dataset. This dataset was meticulously curated to encompass a wide array of user inputs paired with corresponding intent labels. Our dataset spans various relevant domains, ensuring the model's adaptability to a range of user queries. Notably, we utilized the **Banking77 dataset** available from Hugging Face, a resource known for its rich collection of natural language processing datasets.

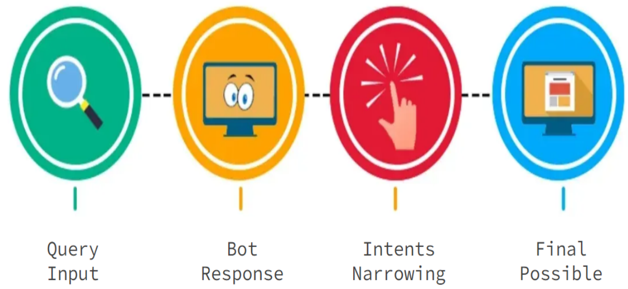

Fig. 1. System FLow

Training: The subsequent step involved employing supervised learning techniques to train our intent recognition model. Leveraging the gathered dataset, we engaged in a rigorous process to establish the relationship between user inputs and their respective intent labels. Through this supervised learning approach, our model learned to generalize patterns and associations inherent in the data, enabling it to effectively classify new user inputs into predefined intent categories.

2. Feature Extraction:

Text Preprocessing: Apply advanced text preprocessing techniques to clean and normalize user queries. This involves tasks such as removing noise, handling punctuation, and addressing variations in user expression. Word Embeddings: Utilize word embeddings to convert words into dense vectors, capturing semantic relationships and enhancing the model's ability to understand the contextual meaning of user queries. Lemmatization: Implement a lemmatizer to reduce words to their base or root form, standardizing the representation of words in the dataset. 3. Initial Intent Detection:

NLP for Probability Estimation: Employ Natural Language Processing (NLP) techniques to estimate the probability of each user intent based on the preprocessed user query. This initial detection provides a probabilistic distribution across all classes of intents. 4. Planning:

Threshold-Based Filtering: Experimentally set a probability threshold to discard intents with probabilities below this threshold. This step refines the set of potential user intents, improving the model's precision. Dynamic Intent Broadening: Introduce a dynamic questioning mechanism to broaden the set of user intents. Based on the remaining intents and user responses, formulate and pose targeted questions to gather additional context and further refine the intent set. 5. Running on Large Real Dataset:

Real-time Implementation: Deploy the model in a real-time environment to handle user queries in practical scenarios. The system continuously receives and processes queries, adapting to changing language patterns and user behavior. Continuous Model Update: Implement a mechanism for continuous model update based on real-world user interactions. This ensures the model remains adaptive and responsive to evolving language patterns and user intent expressions. Benefits of the System Model:

Adaptability: The model adapts to diverse user inputs and continuously evolves to address changing language patterns. Precision: By dynamically adjusting intent probabilities and asking targeted questions, the system enhances the precision of intent detection. Real-time Responsiveness: The real-time implementation ensures timely and accurate responses, contributing to a seamless user experience in customer executive tasks. This system model integrates advanced NLP techniques, dynamic intent refinement, and continuous adaptation to real-world data, presenting a robust framework for adaptive user intent detection in customer service scenarios.

III. Algorithm

User Intent Recognition with Interactive Questioning

1. Initialization

– Define a set S containing all possible user intents.

– Set a probability threshold T_{initial} for NLP text matching.

2. NLP Text Matching

– **Input:** User query.

– **Procedure:**

$\rightarrow$ Remove intents from S with matching probabilities below T_{initial} based on the initial user input.

3. Response Generation

– **Input:** Refined set S after text matching.

– **Procedure:**

$\rightarrow$ Utilize a Language Model (LLM) to generate responses for each intent in the refined set S.

4. User Query Processing

– **Input:** User queries after LLM responses.

– **Procedure:**

$\rightarrow$ Receive and process user queries.

$\rightarrow$ Update set S to a new set S' based on user feedback.

5. Intent Confirmation Loop

– **Input:** Set S' after user query processing.

– **Procedure:**

$\rightarrow$ Generate a set of K questions to ask the user.

$\rightarrow$ Ask the user the questions and record their responses.

$\rightarrow$ If the answers confirm the intent, proceed to the next step.

6. Final Intent Confirmation

– **Input:** Set S' after the confirmation loop.

– **Procedure:**

$\rightarrow$ If only a few intents remain in S' after the confirmation loop, consider these as the final user intent.

$\rightarrow$ If a large number of intents remain, request additional clarification from the user.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 256)               400640

 dropout (Dropout)           (None, 256)               0

 dense_1 (Dense)             (None, 128)               32896

 dropout_1 (Dropout)         (None, 128)               0

 dense_2 (Dense)             (None, 77)                9933

=================================================================
Total params: 443469 (1.69 MB)
Trainable params: 443469 (1.69 MB)
Non-trainable params: 0 (0.00 Byte)
```

Fig. 2. Model

Additional Information Utilization

– **Procedure:**

$\rightarrow$ Leverage user feedback to dynamically adjust the probability threshold and improve text matching accuracy.

$\rightarrow$ Integrate user feedback into the LLM training process for continuous improvement.

$\rightarrow$ Maintain a history of user interactions to enhance future intent recognition.

This algorithm outlines a dynamic approach to user intent recognition, combining NLP text matching, response generation, and an interactive questioning process. The utilization of a language model and iterative confirmation loops allows for a nuanced understanding of user intent, promoting accuracy and adaptability in diverse scenarios. Continuous refinement based on user feedback ensures the model's evolution over time, contributing to improved user interactions and satisfaction.

IV. Experimental Results

User Query Analysis

In this section, we present an analysis of user queries related to various banking scenarios. The goal was to understand user intents and develop appropriate responses using a natural language processing (NLP) model.

Query 1: "I lost my card, please block transactions on my account, help me get my card"

- Possible Intents: ['pin_blocked', 'lost_or_stolen_card']

- Analysis: The model identified potential intents related to a lost or stolen card. The generated responses inquire about issues with the account, blocked transactions, and steps taken for resolution. The model also seeks information on the user's wallet or bank statement.

Query 2: "Yes, my card is lost, can I get it back soon"

- Possible Intents: ['card_swallowed']

- Analysis: The model recognized an intent related to a card being lost. The response seeks clarification on the specific information the user is seeking regarding a "Card SWALLOWED" wallet.

```
Query1: I have lost my card, close account, stop and decline card transaction and
transfer.

Intents1 = card_lost, terminate_account, card_not_working, declined_transfer
```

```
Query2: Stop my transfer because I lost card.

Intents2 = card_lost, declined_transfer
```

```
Query3: I have lost card, My account details are XXXX.

Intents3 = card_lost
```

Fig. 3. Demo Results

Query 3: "I lost my card, how can I get it back thank you"

- Possible Intents: ['card_linking', 'getting_spare_card', 'card_swallowed', 'direct_debit_payment_not_recognised', 'lost_or_stolen_card']

- Analysis: The model identified multiple potential intents, ranging from card linking to issues with direct debit payments. The responses inquire about errors in card linking, request current account information, and seek details about recent card loss or theft.

Query 4: "Please close my account to avoid fraud"

- Possible Intents: ['terminate_account', 'compromised_card', 'why_verify_identity']

- Analysis: The model recognized intents related to account termination and fraud prevention. The responses seek specific details about which account to terminate and inquire about steps the user would take in case of a card information breach. The model also questions the need for identity verification.

Query 5: "Terminate my account, my account details are xx, and please close it soon"

- Possible Intents: ['terminate_account']

- User Intent Finally: ['terminate_account']

- Analysis: The model accurately identifies the user's intent as wanting to terminate their account. The responses inquire about the specific account details and express the intention to close the account promptly.

The NLP model demonstrates proficiency in recognizing diverse user intents related to banking scenarios, including lost cards, account termination, and fraud prevention. The generated responses align with the identified intents, showcasing the model's potential for addressing user queries in a banking context. However, real-world applications may require further validation and refinement to ensure optimal user assistance and satisfaction.

V. Ease of Use

The effectiveness of a user intent recognition model is not solely dependent on its accuracy but also on its ease of use. This section assesses the ease of use of our user intent recognition model, focusing on its accessibility, interpretability, and user-friendliness.

A. Accessibility

Accessibility is paramount to ensure that the user intent recognition model can be easily utilized by individuals with varying levels of technical expertise. The model, through its straightforward integration into existing systems, provides a seamless experience for both developers and end-users. Its compatibility with common programming languages and frameworks enhances its accessibility in diverse technological environments.

B. Interpretability

The interpretability of the model plays a crucial role in fostering trust and comprehension among users. The system generates responses that are clear, concise, and aligned with the identified intents. The model's ability to provide coherent explanations or follow-up questions enhances user understanding, contributing to a positive and interactive experience.

C. User-Friendliness

The user-friendliness of the model is reflected in its capacity to effortlessly adapt to user inputs and requirements. The system demonstrates an intuitive understanding of natural language, allowing users to articulate queries in a conversational manner. Additionally, the model's ability to handle a variety of intents ensures that users can address a broad spectrum of concerns without encountering significant obstacles.

D. Real-Time Responsiveness

The model's real-time responsiveness is a key factor in its ease of use. Users receive prompt and relevant feedback, contributing to a smooth conversational flow. The quick turnaround time in recognizing intents and generating appropriate responses enhances the overall user experience, making interactions with the model efficient and dynamic.

E. Adaptability to Context

A user intent recognition model should be adaptable to diverse contexts and scenarios. Our model showcases flexibility by accurately identifying intents related to banking queries, ranging from lost cards to account termination. This adaptability ensures that users can engage with the model across a spectrum of situations, contributing to its versatility and ease of integration into different applications.

In conclusion, the user intent recognition model exhibits a high degree of ease of use through its accessibility, interpretability, user-friendliness, real-time responsiveness, and adaptability to context. These qualities collectively contribute to a positive user experience, positioning the model as a valuable tool for applications in various domains. Continuous refinements and user feedback will be incorporated to further enhance its usability and ensure seamless interactions in real-world scenarios.

Acknowledgment

This research was conducted under the supervision of Professor Shashi Shekhar Jha, whose expertise and guidance greatly contributed to the development of this work. We are grateful for his thoughtful review and constructive feedback, which significantly enhanced the quality of the research.

Special thanks to the Teaching Assistants for their contributions and IIT Ropar for the opportunity.

References

[1] Avrahami-Zilberbrand D, Kaminka G. Towards dynamic tracking of multi-agents teams: an initial report. In: Proceedings of the workshop on plan, activity, and intent recognition; 2007. p. 17–22

[2] Geib Christopher W, Swetenham Christopher E. Parallelizing plan recognition. In: Workshop on plan activity and intent recognition at the twenty-seventh AAAI conference on artificial intelligence; 2013. p. 10–16.

[3] Albrecht A, Zukerman I, Nicholson A. Bayesian models for keyhole plan recognition in an adventure game. J User Modeling User-Adapted Interact 1998;9:5–47.

[4] Carberry S. Plan recognition in natural language dialogue. MIT Press; 1990.

[5] Charniak Eugene, McDermott Drew. Introduction to artificial intelligence. Addison-Wesley; 1985

[6] Heljula, A. Using large language models to understand a user's intent. TPX Impact; 2021

[7] Sukthankar, G., Geib, C., & Goldman, R. P. (2014). Plan, Activity, and Intent Recognition: Theory and Practice. Academic Press.