# LAPT: Lingua Al Persona Teacher

Yoojin Hwang Yonsei University Seoul, Korea yujinhwang00@yonsei.ac.kr

Maryam Hemati Yonsei University Seoul, Korea hematimaryam@yonsei.ac.kr

SoonChan Kwon Yonsei University Seoul, Korea elrnjs13@yonsei.ac.kr 60 61

66

67

68

69

75

80

81

86

94

96

97

100

101

102

103

105

106

107

108

109

110

111

112

113

114

115

116

DongHee Lee Yonsei University Seoul, Korea dlehdgml1031@yonsei.ac.kr



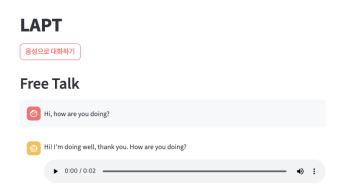


Figure 1: Prototype of our model

#### **Abstract**

10

11

18

19

20

21

22

23

24

25

27 28

34 35

36

37

38

42

43

44

45

47

48

49

50

51

52

54

55

56

57

58

The rapid advancement of artificial intelligence (AI) and natural language processing (NLP) technologies has revolutionized language acquisition process. This paper introduces LAPT (Lingua AI Persona Teacher), an innovative AI-driven tool designed to enhance English-speaking skills through interactive and personalized learning experiences. By leveraging large language models like GPT-4, LAPT provides dynamic, real-time feedback and adaptive learning tailored to individual needs. In addition, considering consumer persona and on-offline touch points, LAPT suggests a variety of contents and gamified challenges to motivate learners. This study outlines the system's features, implementation, and user evaluation, demonstrating LAPT's potential to significantly improve language

# Unpublished working draft. Not for distribution.

needs.

# **Keywords**

Large Language Models, Education, Customization, User-Interaction, Speech-To-Text

proficiency and engagement. Future work will focus on refining

the system and expanding its capabilities to meet diverse learner

#### **ACM Reference Format:**

Yoojin Hwang, SoonChan Kwon, Maryam Hemati, and DongHee Lee. 2018. LAPT: Lingua AI Persona Teacher. In Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 

#### 1 Introduction

The rapid advancement of artificial intelligence (AI) and natural language processing (NLP) technologies has opened up new possibilities for language learning. Among these, AI-driven language models like GPT-4 have demonstrated significant potential in enhancing educational tools by providing interactive and personalized learning experiences. This paper introduces LAPT (Language AI Persona Teacher), an innovative AI-based English-speaking learning tool designed to emulate real-life conversational practice. By

leveraging cutting-edge NLP technologies, LAPT aims to address common challenges in language learning, such as the lack of interaction with native speakers and the need for personalized feedback.

Traditional computer-assisted language learning systems often fall short in providing the dynamic and interactive experiences necessary for mastering a new language. LAPT addresses these short-comings by utilizing large language models to create a more natural and engaging learning environment. This system is designed to help users practice English speaking skills through various interactive modes, including open-ended conversations, structured lessons, scenario-based interactions, and gamified challenges. The main goal of LAPT is to provide a comprehensive and adaptive learning experience that caters to individual learners' needs and preferences.

### 2 Related Works

# 2.1 Large Language Model

A language model (LM) is a model that generates the next word with the highest probability given a context [1]. Recently, with the advent of Transformer models [11], language models have rapidly advanced and impacted many fields. Language models aim to handle complex linguistic features effectively and achieve superior performance in various natural language processing (NLP) tasks by increasing the size of the models, leading to the emergence of a new field called Large Language Models (LLM) [13]. These large-scale language models have significantly influenced the field of natural language processing. Large language models like ChatGPT, Llama2, PaLM2, and GPT-4 not only perform various tasks such as document summarization, translation, and question answering, but they also provide answers to complex questions in different domains. Moreover, they enhance efficiency in daily tasks and activities, such as content creation, programming code analysis, and writing.

# 2.2 Large Language Model for English Education

The field of language learning has significantly progressed with the introduction of LLMs. Traditional computer-assisted language learning systems offered lessons in various subjects but were limited by the need for learners to interact with real people, like teachers or peers, for practicing actual conversations [9]. LLMs have the potential to address these limitations by providing more natural interactions and feedback that closely mimic human responses [8].

Owing to these advantages, recent studies have explored the application of Large Language Models (LLMs) in English education. For instance, one study proposed a model that estimates individual students' knowledge states and generates adaptive, personalized practice problems accordingly [2]. Another study introduced PEEP-Talk, a conversation-based chatbot designed for English education that offers topic transitions and grammar feedback [9]. Additionally, research has suggested methods for creating practice problems tailored to learning objectives by automatically identifying and generating distractors through dynamic learning analysis and error diagnosis modules [4]. These investigations demonstrate the potential of LLMs in enhancing English language education.

#### 2.3 Personalization

The review on personalization in user interfaces (UI) and user experiences (UX) is categorized into two primary areas: conceptual frameworks / classifications and implementation / evaluation. The first group includes papers that construct a comprehensive conceptual framework through extensive literature reviews and propose detailed classification schemes for personalization [12] [3]. The second group focuses on implementing and evaluating personalized UI/UX systems, addressing challenges like uniform UI designs and limited customization through systematic methodologies [6], and developing intelligent UI/UX systems tailored for elderly users utilizing clustering techniques and classifiers [10].

Further advancements in this field address static UI behavior in recommendation systems [5] and optimize UI widget assembly using contextual Multi-Armed Bandit (MAB) algorithms [7].

Personalization in various disciplines encompasses distinct definitions. In marketing and e-commerce, it tailors interactions to build one-to-one relationships [3]. Cognitive science views it as making assumptions about users' goals and preferences based on behavior. In computer science, personalization involves technologies designed to accommodate individual user patterns using machinelearning algorithms. Additionally, personalization affects spatial dimensions in architectural psychology and prioritizes relevant information in information science.

Research on personalized UI/UX design integrates cognitive and behavioral models to tailor interfaces, validated through usability and satisfaction evaluations [6]. The SmartSenior system optimizes interfaces for senior citizens using clustering techniques [10]. In recommendation systems, personalized UI evolves through continuous user feedback and contextual information [5]. A framework using MAB algorithms addresses dynamic and contextual challenges in UI widget development [7].

Identified gaps and future research recommendations include enhancing understanding of personalization's impact on performance, fostering interdisciplinary collaboration, refining user models with diverse user bases, broadening research factors, and improving clustering accuracy. Online cognitive testing, increasing sample size, standardized usability testing, exploring objective proficiency measures, and addressing MAB algorithm challenges at scale are also suggested [5–7, 10].

These studies highlight the effectiveness of personalization in UI/UX design and service performance, suggesting a process of iterative refinement based on user feedback and data clustering to achieve effective personalized designs.

### 3 User Study

In South Korea, English education encompasses a wide and diverse range of targets, educational scopes, and objectives. In exploring the potential for a new market, a User Study was conducted to identify the pain points of the existing market and to understand the needs of consumers.

The User Study focused on five key attributes essential for English education. Five consumers from various demographic backgrounds were randomly selected for this study. The attributes and results are as follows:

- (Attribute 1) Fluency in English Use

233 -234 -235 -236 -

- (Attribute 2) Frequency of English Use in Daily Life
- (Attribute 3) Importance in English Education
- (Attribute 4) Acceptance of New Technology
- (Attribute 5) Availability of Time in Daily Life

Among the overall survey participants, consumers with similar user attributes were categorized into individual personas. Each persona was defined as 'Creative Alice' and 'Strict Steve'. The attribute results for the two personas are presented in the table below.

Table 1: Comparison of Attributes of Each Persona

|             | Alice | Steve |
|-------------|-------|-------|
| Attribute 1 | Low   | High  |
| Attribute 2 | Low   | High  |
| Attribute 3 | Mid   | High  |
| Attribute 4 | Mid   | Mid   |
| Attribute 5 | High  | Low   |

Creative Alice is a mid-20s female marketer with an introverted personality who has limited experience with English education for conversational purposes outside the formal education system. Although it has been a long time since she last studied a foreign language, she has planned to learn a new language when needed. Due to her unsocial nature, she prefers learning English through mobile applications rather than conversing with real people.

In contrast, the persona of Strict Steve is an extroverted male in his late 20s. He is a strict leader-like office worker with excellent English proficiency, frequently using English in his daily life. Furthermore, having already acquired a certain level of English proficiency, he seeks an efficient learning method that accommodates his regular schedule to achieve a higher level of English skill.

Based on the extensive range of personas and needs revealed by the User Study results, a CEJ (Customer Experience Journey) Map was additionally created. This map visualizes the experience journey of consumers using English education services, aiming to explore the touchpoints with users at each stage and identify improvement opportunities.

The stages of the CEJ are broadly composed of Research, Purchase Decision, Use, and Feedback & Review, which can be explained as follows:

- Research: Touch points can be TV/Radio commercial, social media contents (Instagram/Youtube), outdoor Ad. So, they could be provided of creative advertisement contents about the service, like a unique catch phrase.
- Purchase Decision: Touch points can be recommendation system on website or discount coupons. So, it would be a great idea to provide discount promotion, send pop-up messages and compare current discounted price with other services/customers.
- Use: Touch points are motivating consistent training through pop-up notice/alarm, sending useful information about language acquisition, collecting diverse VOC for new/existing users, developing qualified and trendy education contents and releasing creative curation contents.

Feedback and Review: Touch points are VOC center on the
website and reviewing about marketing campaign on social
media. So the service can give additional promotion if they
got recommendation from original users and build customer
segmentation data for new marketing strategies at the same
time.

Specifically, during the CEJ research stage, a brief case study was conducted on the new service (LAPT) presented in this study as a benchmark-worthy example. In the Research stage, Speak was investigated as a best practice example for enhancing brand awareness among consumers through creative advertising. In the Use stage, Toss Feed was examined as an example of providing creative curated content.

# 4 System Design

Our project involves the development of an AI-based Englishspeaking learning tool, akin to applications like Speak. The system is designed to facilitate interactive and immersive language learning experiences using advanced natural language processing techniques. The primary objective is to create an engaging platform where users can practice English speaking, receive real-time feedback, and improve their language skills through conversational interactions.

# 4.1 Features and Contents

The design of our application was meticulously crafted using Figma, a powerful design tool that enabled us to create a seamless and user-friendly interface. The mobile design of our application incorporates several key features aimed at enhancing the user experience and ensuring effective learning:

#### 4.1.1 Features.

- Interactive User Interface: The interface is designed to be intuitive, with easy-to-navigate menus and clear prompts to guide users through their learning journey. Each screen is optimized for mobile devices to ensure a smooth user experience. Users can easily access different features such as starting a conversation, reviewing conversation history, and settings.
- Voice Interaction: Users can interact with the system using voice commands, making the learning process more natural and similar to real-life conversations. This feature allows users to practice speaking in a way that mirrors actual dialogue, helping them to improve their fluency and confidence. The voice interaction capability is designed to be highly responsive and accurate, providing a seamless conversational experience.
- Real-time Feedback: The application provides immediate feedback on pronunciation, grammar, and fluency, helping users to correct mistakes and improve their speaking skills promptly. This feature leverages advanced speech recognition and natural language processing technologies to deliver precise and constructive feedback.
- Adaptive Learning: The system adjusts the difficulty and content of the conversations based on the user's proficiency level,



Figure 2: Main Contents.

ensuring a personalized learning experience. This adaptability is achieved through Large Language Model that analyze user performance and modify content dynamically. The adaptive learning feature ensures that users are always challenged appropriately and can progress at their own pace.

• Gamification: Elements such as badges, points, and progress tracking are incorporated to motivate users and make learning more enjoyable. Gamification strategies are implemented to encourage regular practice and reward milestones. Users earn badges and points for completing tasks and reaching milestones, which helps to maintain motivation and engagement. The gamification elements are designed to be visually appealing and rewarding, adding an element of fun to the learning process.

Utilizing these features, our application offers several distinct contents that provide varied and engaging learning experiences. These contents are designed to cater to different learning preferences and goals, ensuring that users have access to a comprehensive suite of tools for improving their English speaking skills. An image of the contents can be seen in Figure 1.

#### 4.1.2 Main Contents.

- Free Talk: This is the most basic function of the application, allowing users to engage in open-ended conversations. Users can practice their speaking skills freely on any topic they choose, providing a versatile and flexible learning experience. All conversations in Free Talk are saved and LAPT automatically summarizes the overall contents of each conversation. In addition, users can practice various expressions learned throughout the lessons of LAPT whenever they want. Finally, there are diverse characters with its own unique personality, so that the character can customize to the personality of a user.
- Situation: In this content, users engage in conversations based on specific scenarios, such as ordering food at a restaurant or checking into a hotel. The Situation helps users to practice practical language skills and prepare for real-life interactions. It enhances the user's ability to handle various situations in English-speaking environments.
- Learn with Celebrities: This content presents videos featuring various celebrities and movie characters using authentic conversational English expressions, accompanied by subtitles, to help users directly acquire their vivid expressions. This approach is based on the idea that learning engagement can be significantly enhanced if users can learn from their favorite celebrities. This unique feature, not found in other existing services, is proposed as a distinctive point of differentiation for LAPT.
- Challenges: This content includes various activities that challenge users to use their English skills in fun and competitive ways. Challenges are related to the gamification aspects, designed to test and improve different aspects of language proficiency. The Challenges helps to reinforce learning through play, making the process both effective and enjoyable. The missions in the challenge are categorized into daily and monthly missions. The number of days a user completes the daily missions is counted and reflected in the monthly challenge achievement. Users can enhance their motivation by checking their real-time ranking among all users.

## 4.1.3 Sub Contents.

• My Page: The "My Page" section of LAPT is meticulously designed to enhance user convenience and accessibility, even beyond the platform's core functionalities. The integration of a calendar feature facilitates the addition of events related to LAPT, such as subscription renewals, challenge deadlines, and anniversaries. Additionally, the New Reminder function allows users to freely incorporate their personal schedules, thereby offering greater flexibility. The section consolidates features like Bookmarks, My Favorite, and Watch Later to support uninterrupted learning experiences. Moreover, the History and Report features provide comprehensive summaries of the user's learning activities regularly, enabling users to track their progress effectively.



Figure 3: My Page.

- Report: Each user can freely set specific periods and select items to be reported about their performance and learning process, providing a learning report similar to a report card. The extent of implementation of LAPT's core services is detailed, especially for users who utilize voice services in Free Talk. In detail, the Avatar AI checks pronunciation and presents scores in a detailed manner. For instance, out of 1,234 words, 990 are pronounced correctly, 200 have minor spelling errors, and 44 are pronounced incorrectly.
- Search: The "Search" page is designed to allow users to explore desired content at any time. To maximize user retention, example search phrases are provided to assist users who may have forgotten what they want to search for, clicked the wrong button, or clicked unintentionally. Features such as recent searches and recommended search terms help remind users of relevant content. Additionally, trending search terms explain the expressions and content that other users are currently interested in, fostering a more engaging and dynamic user experience.

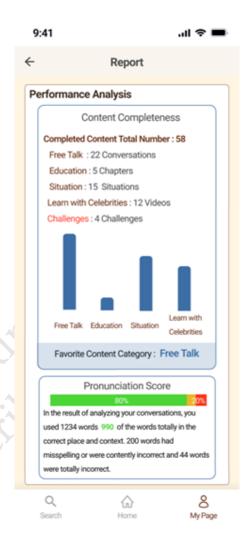


Figure 4: Report.

# 4.2 Implementation

The implementation of our system leverages several cutting-edge technologies and libraries to ensure robust performance and scalability. The code for the system is primarily written in Python, utilizing the following key components: Streamlit, a framework for creating interactive web applications, used to develop the frontend interface and manage user interactions, making it easy to deploy and iterate on the application; OpenAI API, integrated to utilize advanced language models such as GPT-4 for generating responses and conducting natural language processing tasks, enabling the system to deliver human-like conversational experiences; SpeechRecognition, a Python library for performing speech recognition, used to capture and process voice inputs from users, converting spoken language into text for further processing; pydub, a library for manipulating audio files, used to handle audio data and convert it into formats suitable for processing by other components, ensuring high-quality audio input and output; Microphone and AudioSegment, components used to capture and process audio inputs

from the users, ensuring that the audio data is clear and suitable for analysis by the speech recognition system.

Our system captures audio input from the user using a microphone, adjusts for ambient noise, and saves the recorded audio in WAV format. The recorded speech is then converted into text using the OpenAI API. This step is vital for understanding the user's input and generating appropriate responses. The recognized text is stored in the session state for further processing, ensuring that the conversation flow remains coherent and contextually relevant.

To provide a natural conversational experience, the system converts generated text responses back into speech. This bidirectional interaction mimics real-life conversations and enhances the learning experience. Users receive audio feedback, which helps them improve their listening and speaking skills simultaneously.

The main interface is created using Streamlit, where users can start a conversation by pressing a button that triggers the speech-to-text process. The recognized text is used as a prompt for generating responses using GPT-4, and the interaction history is displayed to the user. This interactive loop ensures that users receive immediate feedback and can engage in continuous practice.

The architecture of the system is designed to support real-time interaction and scalable performance. Although still in the prototype stage, we believe that with further development, we can provide a comprehensive solution to effectively learn English speaking skills interactively by utilizing advanced AI models and a robust software framework. We also believe that by using cutting-edge technology to provide users with accurate feedback and a natural conversational experience, we will greatly enhance the language learning process.

#### 5 User Test

After constructing a low-fidelity system design, a System Usability Scale (SUS) evaluation was conducted to assess basic usability. The evaluation participants were different from those in the User Study, and they were randomly selected in terms of age and gender to minimize bias within the sample.

The evaluation consisted of 10 questions assessing overall usability, and responses were collected on a 5-point Likert scale, resulting in a total score of 79.25.

The main feedback from users indicated that the app has too many features, making it somewhat difficult to learn how to use according to the main stream. However, they also noted that the system's utility is high and the integration of features is well-regarded.

To receive more detailed feedback and suggestions for improvement, subjective interviews were also conducted alongside the SUS. The interview questions focused on overall impressions of the service, strengths and weaknesses, and suggestions for improvement. The main responses and their incorporation into the high-fidelity design are as follows:

#### [Feedback]

- "The hierarchy of information needs to be better reflected in the design."
- "The font size appears unnecessarily large."
- "The content to be learned today and the lesson to be performed should be more prominently designed at the beginning."

- "Add explanations about the chatbot's personality on the chatbot selection page."
- "Provide more diverse content tailored to individual preferences."
- "A strategy is needed to guide users through the app's functions"
- "The connection between functions feels somewhat low."
- "The challenges need more variation to enable consistent participation."
- "It would be better if the script or subtitles in the videos were more visible."

#### [Results]

- "Reflect the importance of LAPT's key feature, Free Talk, by representing it in a large box at the top."
- "Standardize and reduce the overall font size."
- "Support the display of learning status with Active Level indicators on the main screen."
- "Enable explanations and selection of AI character MBTI in Free Talk."
- "Activate the Customization Setting feature."
- "Add a 'Browse Pages' feature on the first screen."
- "Add functionality to immediately use learned expressions in Free Talk."
- "Add a user-specific challenge ranking screen."
- "Provide the script within videos and add functionality to move to the screen when clicked."

### 6 Discussion

The development of LAPT represents a significant step forward in the application of AI technologies in language education. One of the most notable features of LAPT is its ability to provide real-time feedback on pronunciation, grammar, and fluency. This immediate feedback is crucial for language learners, as it allows them to identify and correct mistakes as they occur, leading to more effective learning outcomes. The use of advanced speech recognition technology ensures that the feedback is accurate and relevant, further enhancing the learning experience.

Another key aspect of LAPT is its adaptive learning capabilities. By analyzing user performance and adjusting the difficulty and content of conversations accordingly, LAPT ensures that learners are always appropriately challenged. This personalized approach helps maintain learner engagement and motivation, as users can see their progress and are consistently pushed to improve their skills. The incorporation of gamification elements, such as badges and points, adds an additional layer of motivation, making the learning process more enjoyable and rewarding.

The diverse content modes offered by LAPT, including Free Talk, Education, Situation, Learn with Celebrities, and Challenges, provide a well-rounded approach to language learning. Each mode caters to different aspects of language acquisition, from practicing spontaneous conversation to mastering specific vocabulary and grammar. This variety ensures that learners can find the type of practice that best suits their needs, further enhancing the effectiveness of the tool.

Despite its many strengths, LAPT is not without its limitations. As a prototype, the system still requires further development and

756

757

760

761

762

766

767

768

769

770

771

772

773

774

781

782

787

793

794

795

796

797

800

801

802

807

808

809

810

811

812

749

750

751

752

753

754

testing to fully realize its potential. Future work will focus on refining the speech recognition and feedback mechanisms, expanding the range of conversational scenarios, and improving the overall user experience. Additionally, more extensive user testing will be conducted to gather feedback and identify areas for improvement.

# Conclusion

In conclusion, LAPT (Lingua AI Persona Teacher) represents a promising advancement in AI-driven language learning tools. By leveraging the capabilities of large language models and advanced speech recognition technology, LAPT provides a dynamic and interactive platform for practicing English speaking skills. The system's e, re real-time feedback, adaptive learning capabilities, and diverse content modes make it a comprehensive solution for language learners seeking to improve their proficiency. While still in the prototype stage, LAPT has demonstrated significant potential in enhancing the language learning process and addressing common challenges faced by learners. Future development and testing will further refine the system, ensuring that it continues to meet the needs of users and remains at the forefront of AI-driven educational tools.

By combining cutting-edge AI technologies with a user-centered design, LAPT offers a novel approach to language learning that is both effective and engaging. As the system continues to evolve, it has the potential to significantly impact the field of language education, providing learners with the tools they need to achieve fluency and confidence in their English speaking skills.

# References

- [1] Yoshua Bengio, Réjean Ducharme, and Pascal Vincent. 2000. A neural probabilistic language model. Advances in neural information processing systems 13 (2000).
- [2] Peng Cui and Mrinmaya Sachan. 2023. Adaptive and personalized exercise eneration for online language learning. arXiv preprint arXiv:2306.02457 (2023).
- [3] Haiyan Fan and Marshall Scott Poole. 2006. What is personalization? Perspectives on the design and implementation of personalization in information systems. Journal of Organizational Computing and Electronic Commerce 16, 3-4, 179–202.
- Tanja Heck and Detmar Meurers. 2023. Using Learning Analytics for Adaptive Exercise Generation. In Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023). 44-56
- Jamil Hussain, Wajahat Ali Khan, Muhammad Afzal, Maqbool Hussain, Byeong Ho Kang, and Sungyoung Lee. 2014. Adaptive user interface and user experience based authoring tool for recommendation systems. In Ubiquitous Computing and Ambient Intelligence. Personalisation and User Adapted Services: 8th International Conference, UCAmI 2014, Belfast, UK, December 2-5, 2014. Proceedings 8. Springer, 136-142.
- Hyesung Ji, Youdong Yun, Seolhwa Lee, Kuekyeng Kim, and Heuiseok Lim. 2018. An adaptable UI/UX considering user's cognitive and behavior information in distributed environment. Cluster Computing 21, 1045-1058.
- [7] Ioannis Kangas, Maud Schwoerer, and Lucas J Bernardi. 2021. Recommender systems for personalized user experience: lessons learned at Booking, com. In Proceedings of the 15th ACM Conference on Recommender Systems. 583-586.
- Na-Young Kim, Yoonjung Cha, and Hea-Suk Kim. 2019. Future English learning: Chatbots and artificial intelligence. Multimedia-Assisted Language Learning 22, 3
- Seungjun Lee, Yoonna Jang, Chanjun Park, Jungseob Lee, Jaehyung Seo, Hyeonseok Moon, Sugyeong Eo, Seounghoon Lee, Bernardo Yahya, and Heui-Seok Lim. 2023. PEEP-Talk: A Situational Dialogue-based Chatbot for English Education. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations). 190-207.
- [10] Heuiseok Lim, Danial Hooshyar, Hyesung Ji, Seolhwa Lee, and Jaechoon Jo. 2019. SmartSenior: automatic content personalization through semi-supervised learning. Wireless Personal Communications 105, 461-473.
- [11] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. Advances in neural information processing systems 30 (2017).
- Jari Vesanen. 2007. What is personalization? A conceptual framework. European Journal of Marketing 41, 5/6, 409-418.

[13] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al. 2023. A survey of large language models. arXiv preprint arXiv:2303.18223 (2023).

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009

2024-06-22 14:03. Page 7 of 1-7.