Enhancing Intelligent Voice Assistants in Tourism Services A Computational Approach

Zhiyu Zhang*

Industrial Development Research Center of Sichuan University, Chengdu, China
zyzhang0603@163.com

Abstract. This study aims to enhance the response efficiency of intelligent voice assistants in tourism industry services by analyzing technical performance improvements and application strategies. It thoroughly examines the application of these assistants in tourism, assessing their functionality in information retrieval, travel booking, payment services, navigation explanations, cross-language translation, customer opinion collection, and satisfaction evaluation. The paper then showcases how upgrading speech recognition and natural language processing technologies can optimize processing strategies and personalize services. Employing advanced computational models, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), the study enhances the assistants' accuracy in noisy environments and multilingual natural language understanding. Additionally, large-scale multilingual pre-trained models like mBERT and mT5 are used to boost cross-cultural adaptability and language translation capabilities. By integrating these technologies, the research demonstrates significant enhancements in operational efficiency and user experience, thereby providing the tourism industry with more efficient and personalized service solutions through improved adaptability in complex linguistic and cultural environments.

Keywords: intelligent voice assistant, travel service, speech recognition, natural language processing

Introduction

Against the backdrop of rapid advances in artificial intelligence technology, intelligent voice assistants are increasingly being used in the tourism service industry, which has greatly improved the efficiency and customer satisfaction in information retrieval, reservation services, tour guide interpretation, and other aspects. However, at this stage, intelligent voice assistants still show certain limitations in terms of accuracy of speech recognition, depth of natural language understanding, and provision of personalized services, which restricts further improvement of their service responsiveness. This study is dedicated to exploring the application of intelligent voice assistants in the tourism service industry, and based on this, proposes strategies to enhance speech recognition and natural language processing technology, improve personalized services, promote multilingual and cross-cultural adaptability, and strengthen security and privacy protection. The purpose is to enhance the interactive responsiveness of intelligent voice

M. R. Mohyuddin and N. A. D. IDE (eds.), *Proceeding of the 2024 International Conference on Diversified Education and Social Development (DESD 2024)*, Advances in Social Science, Education and Humanities Research 899,
https://doi.org/10.2991/978-2-38476-346-7_6

assistants in tourism services and provide reference for the intelligent advancement of the tourism industry.

1. Current application status of intelligent voice assistants in tourism services

Definition and Technical Principles of Intelligent Voice Assistant

Intelligent voice assistant is a technical system that realizes voice interaction with users and provides information and services through key links such as speech recognition, natural language processing, and speech synthesis. As shown in Figure 1, initially, users use voice to issue search instructions. , Subsequently, the system uses speech recognition means to convert the speech signal into text information, and uses natural language processing technology to conduct in-depth analysis of the text content to reveal the real needs of the user. Subsequently, the system connects to the knowledge base or related information through the API interface. Database, performs retrieval functions and processes query information raised by users. Finally, the processed data is converted into voice information through speech synthesis technology, and this information is delivered to the user by the voice assistant. The function of the intelligent voice assistant is not limited to instant information query, it can also perform reservation services, provide tour guide translation, and collect customer feedback. These features significantly improve the efficiency and user experience of the service.

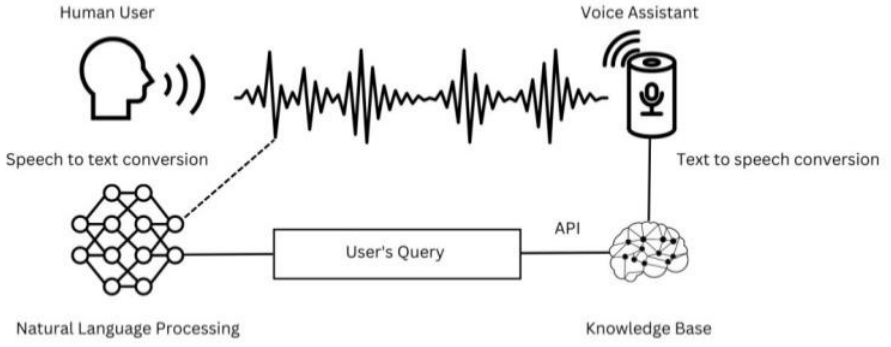

Fig. 1. Working principle diagram of intelligent voice assistant

Application of Intelligent Voice Assistant in Tourism Services

1) Tourism Information Query

In terms of travel information inquiry, intelligent voice assistants can significantly improve the convenience and efficiency of obtaining information. Users only need to input the desired destination, attraction introduction, transportation guide and other information through voice, and the voice assistant can use its voice recognition and natural language processing technology to quickly convert the user's voice into text, analyze the user's intention, and find relevant information in the knowledge base. The user's query is converted into text form through the speech recognition system, and then enters the natural language processing stage. The system will search for matching

travel information based on the keywords entered by the user and generate a detailed answer accordingly. Finally, speech synthesis technology is used to convey the answer to the user again in the form of speech. When the user asks "the opening hours of the Forbidden City in Beijing", the voice assistant will convert the user's voice into text, and then identify the keywords "Forbidden City in Beijing" and "Opening time", and then retrieve relevant information in the knowledge base and feedback the query results to the user in voice form. This process improves the quality of user interaction by optimizing the efficiency of information retrieval, thereby making the retrieval of travel information more intelligent. ization and personalization.

2) Booking and Payment Services

Users can use the intelligent voice assistant to perform various operations such as hotel reservations, flight ticket orders, and restaurant seat reservations. This system uses voice recognition and natural language processing technology to convert the user's voice instructions into executable commands. By calling the corresponding API interface, it interacts with major booking platforms to complete the entire booking process. When the user expresses a certain demand, such as assisting in booking a flight to Shanghai, and the time is set for tomorrow morning, the voice assistant will parse this demand and search the airline's booking system to find flight data that meets the user's requirements. After the user confirms, the system will guide the user to make online payment operations through integrated payment interfaces such as Alipay and WeChat Pay. This process simplifies the user's operations in the booking and payment links, while enhancing the portability and protection of travel services, ensuring that consumers can obtain efficient travel support at any time and any place. The intelligent voice assistant can use the user's past booking records and personal preferences to provide customized suggestions, thereby better improving the user experience.

3) Tour guide and translation services

During the tour, tourists can use the voice assistant as a technical tool to easily obtain detailed explanations and historical information of the scenic spots without relying on traditional written materials. When tourists ask questions about "the history of this building", the voice assistant will convert the tourists' voice into text, and then identify the keywords "architecture" and "history". Then, it will retrieve relevant content from the preset knowledge base and use speech synthesis technology to explain it to tourists in detail. One of the functions of the voice assistant is to provide real-time translation services, which helps tourists who use different languages to overcome language barriers. When the user asks "Where is the nearest toilet?", the voice assistant will translate it into Chinese and provide the location information of the nearest toilet. The application not only enhances the travel experience, but also acts as a medium of communication in the exchange between different cultures, thereby promoting tourism services to become more diversified and personalized.

4) Customer Feedback and Satisfaction Survey

During or after the trip, consumers can submit their feedback through the intelligent voice control function. The system uses voice recognition software to convert voice input into written text, and then uses natural language processing methods to summarize and analyze the text information obtained. For example, in the user experience evaluation, the satisfaction with the tour guide part is higher, and improvement suggestions are made for the catering part. Then, the system will store this information in the database

Companies can use satisfaction scoring formulas $Satisfaction\,Score = \frac{\text{Positive Feedback}}{\text{Total Feedback}} \times 100\%$ to analyze the collected data and understand user satisfaction and improvement needs.

2. Strategies for Intelligent Voice Assistants to Improve the Responsiveness of Tourism Services

Improving speech recognition and natural language processing technologies

1) Improving the accuracy of speech recognition

Improving the accuracy of speech recognition is the key to improving the quality of intelligent voice assistant services. For the recognition problems in noisy environments and accents and dialects, convolutional neural networks (CNN) and recurrent neural networks (RNN) in deep learning algorithms can be used for multi-level feature extraction and sequence modeling. For example, by optimizing the loss function, $\mathrm{L} = -\sum_{i=1}^{N} \log P\left(y_i \mid x_i\right)$ where is the probability of $P\left(y_i \mid x_i\right)$ predicting the correct label y_i when a given input is given x_i, and by increasing the diversity and scale of the data set, the generalization ability of the model is enhanced. In addition, the adaptive noise cancellation (ANC) technology is used. Through a mathematical model $y(t) = x(t) + n(t)$, where $y(t)$ is a noisy speech signal, $x(t)$ is a pure speech signal, and $n(t)$ is a noise signal, the pure speech signal is extracted through filtering and noise reduction algorithms to further improve the accuracy of speech recognition.

2) Enhancing the depth of natural language understanding

Enhancing the depth of natural language understanding is an important step in improving the responsiveness of intelligent voice assistants. Using the attention mechanism and transformer model in deep learning can significantly improve the understanding of complex sentences and the effect of multi-turn dialogue processing. By optimizing the loss function $\mathrm{L} = -\sum_{i=1}^{N} \log P\left(y_i \mid x_i, \text{context}\right)$, where context is context information, and through weighted calculation of context information, the accuracy and coherence of semantic understanding are improved. In addition, the bidirectional long short-term memory network (Bi-LSTM) is used to improve the comprehensive understanding of sentence structure and semantics through forward and backward contextual information. $h_t = \sigma\left(W_h x_t + U_h h_{t-1} + b_h\right)$

Personalized services and user experience optimization

1) Personalized recommendations based on user data

Personalized recommendations based on user data can significantly improve the service quality and user experience of intelligent voice assistants. Build a personalized recommendation model by analyzing users' historical behavioral data and preferences.

For example, the collaborative filtering algorithm is used to find similar users and recommend their favorite items by calculating user similarity based on the user's historical rating matrix $R = \{r_{ui}\}$, which r_{ui} represents user u's rating of item i.

$$sim(u,v) = \frac{\sum_{i \in I}(r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sqrt{\sum_{i \in I}(r_{ui} - \bar{r}_u)^2}\sqrt{\sum_{i \in I}(r_{vi} - \bar{r}_v)^2}}$$

In addition, a content-based recommendation algorithm is used to extract the sum of feature vectors of users and items $\vec{u}$, $\vec{i}$ calculate the similarity $sim(\vec{u},\vec{i}) = \frac{\vec{u} \cdot \vec{i}}{|\vec{u}||\vec{i}|}$, and recommend content that matches the user's interests.

2) Optimize user interaction interface and experience

Optimizing the user interaction interface and experience is an important aspect of improving the practicality of intelligent voice assistants. First, by improving the user interface (UI) design, the efficiency and comfort of interaction between users and voice assistants can be improved. For example, a simple and intuitive interface design and a humanized interaction process can be adopted to enable users to quickly find the required functions and information. Secondly, by applying human-computer interaction (HCI) technology and optimizing the dialogue management strategy, voice assistants can conduct multiple rounds of dialogue and task switching more naturally and smoothly. Reinforcement learning algorithms can be used to define reward functions $R(s,a)$ to evaluate the effect of each interaction action a in state s, and the dialogue strategy can be continuously optimized. In addition, sentiment analysis technology can be used to analyze the emotional characteristics in the user's voice $S = f(\text{audio features})$, adjust the response mode and tone of the voice assistant, and provide more intimate and personalized services.

Multilingual Support and Cross-Cultural Adaptation

1) Expanding multi-language recognition and translation capabilities

By introducing advanced neural machine translation technology, translation quality and speed can be significantly improved. Specifically, a sequence-to-sequence (Seq2Seq) model based on the attention mechanism is used to convert the input speech signal into text in the target language. For example, the loss function in the model can be defined as cross-entropy loss $\mathrm{L} = -\sum_{t=1}^{T} \log P(y_t | y_{<t}, x)$, where x represents the input sequence and y_t represents the t-th word of the target language. By optimizing the loss function, the accuracy of translation is improved. In addition, using multi-language pre-trained models, including mBERT or mT5, can handle multiple languages at the same time and enhance the system's multi-language processing capabilities.

2) Improve the flexibility of cross-cultural interaction

Through the cultural adaptation model, the intelligent voice assistant can adjust the interaction method and content according to the user's cultural background. For exam-

ple, a rule-based cultural adaptation strategy is adopted to define the interaction rule sets of different cultures, and the behavior and response methods of the voice assistant are dynamically adjusted during the interaction process. In addition, the transfer learning technology in machine learning is applied to transfer the knowledge of a specific culture to a new cultural background through training on different cultural data sets. The specific method includes optimizing the transfer loss function $L_{transfer} = L_{source} + \lambda L_{target}$, where L_{source} and L_{target} are the data losses of the source culture and the target culture respectively, and λ is the weight coefficient. By balancing the knowledge transfer of the source culture and the target culture, the performance of the model in the new cultural background is improved.

3. Conclusion

Research in the tourism services sector demonstrates that intelligent voice assistants significantly enhance information retrieval, reservation services, guide translation, and customer feedback mechanisms. To further boost service efficiency, it is crucial to advance current voice recognition technologies and augment the capabilities of natural language processing algorithms. Enhancing the personalization of recommendation systems and refining user interface designs are also imperative. Moreover, expanding multilingual support and bolstering cross-cultural adaptability are essential to elevate the comprehensive application of intelligent voice assistants in the tourism industry. Technological advancements, particularly in computer science, big data analytics, and artificial intelligence, will enable these assistants to play an increasingly pivotal role. As these technologies evolve, intelligent voice assistants are poised to offer more convenient, efficient, and tailored service experiences, thus driving the intelligentization and modernization of the tourism industry.

References

1. Du Jiangang, Zhao Huan, Su Jiuru, et al. Customer behavior under service intelligence: Research review and prospects[J]. Foreign Economics and Management, 2022, 44(3).
2. Xu Xiaoyan. On the application of intelligent voice interaction technology in hotels[J]. Electronic Communications and Computer Science, 2021, 3(5).
3. Liu Xin, Xie Lishan, Li Dongmei. Research on the impact of tourism service robot anthropomorphism on customer value co-creation willingness[J]. Tourism Tribune/Lvyou Xuekan, 2021, 36(6).
4. Liu Chun, Lu Ximing, Zhao Qi. A review of research on contactless services in tourism and hotel industry[J]. Tourism Guide, 2024, 8(1): 104.
5. Chai Runze, Sun Mingyang. Research on the design of car voice assistant based on human-computer interaction technology[J]. Industrial Design, 2021 (1): 30-31.