Article

Not peer-reviewed version

Deep Learning-Enhanced Adaptive Interface for Improved Accessibility in E-Government Platforms

XIANG LI and Yikan Wang *

Posted Date: 12 December 2024

doi: 10.20944/preprints202412.1062.v1

Keywords: Deep learning; Adaptive interfaces; E-government; Low literacy rates; Special needs

doi:10.20944/preprints202412.1062.v1

Article

Deep Learning-Enhanced Adaptive Interface for Improved Accessibility in E-Government Platforms

Xiang Li [1] and Yikan Wang [2,*]

[1] The School of Visual and Performing Arts, Long Island University Post, Brookville, NY, 11548, USA
[2] School of Systems and Enterprises, Stevens Institute of Technology, New Jersey, 07030, USA
* Correspondence: ywang463@stevens.edu

Abstract: The advent of e-government platforms has highlighted the pressing need to enhance the accessibility of digital services for individuals with low literacy rates and special needs. The objective of this study is to design and develop an adaptive interface system utilising deep learning technology with the intention of markedly enhancing the accessibility and user-friendliness of e-government platforms, particularly for users with low literacy rates and special needs. The proposed model incorporates multimodal learning and reinforcement learning mechanisms, enabling automatic adjustments to the interface layout, text size, colour contrast and voice assistance functions to align with the personalised requirements of users. In particular, when integrated with sentiment analysis and speech recognition modules, the system is capable of identifying and predicting the user's operational intent in real time, thereby enabling dynamic interface adjustments. The model's innovative aspect lies in its adaptive adjustment capability and high efficiency in low-resource environments. The experimental results demonstrate that the proposed model can markedly enhance the operational efficiency and satisfaction of individuals with low literacy rates and special needs, while also improving the user experience.

Keywords: deep learning; adaptive interfaces; E-government; low literacy rates; special needs

1. Introduction

The advent of the digital age has seen the e-government platform emerge as a pivotal component of contemporary public services, offering a seamless and effective interactive platform for government and citizen interaction. However, the majority of these platforms exhibit shortcomings in their design, particularly with regard to accessibility for demographic groups with low literacy rates and specific needs. Globally, despite the introduction of relevant policies in numerous countries and regions to advance the development of e-government, the design of the platform continues to inadequately address the specific requirements of these groups. As reported by the United Nations, approximately 30% of the global population has some form of disability, and the majority of individuals reside in regions with low literacy rates, which encounter numerous challenges in their daily lives [1]. For these groups, the standardised and singular interface design of e-government platforms often proves inadequate, resulting in significant difficulties in their use.

Therefore, enhancing the accessibility of e-government platforms has become a pressing challenge. The conventional methodology for accessibility design predominantly concentrates on visual and auditory stimuli as the principal inputs, neglecting the necessity for personalisation and dynamic adaptation. In recent years, there have been notable advances in deep learning technology, particularly in the domains of image processing, speech recognition, and natural language processing. These developments have led to the emergence of novel approaches to enhancing accessibility. The use of deep learning enables the interface to be adaptively adjusted in real time according to the needs of the user [2]. This allows for the dynamic optimisation of visual elements, such as interface layout, text size and colour contrast, as well as the provision of more personalised and intelligent auxiliary functions in conjunction with input methods, such as voice and gestures.

 doi:10.20944/preprints202412.1062.v1

This has the potential to enhance the user experience of groups with low literacy rates and special needs [3].

In this process, the comprehensive promotion of government affairs informatisation, particularly the construction of an integrated government service platform, has become the core task of governments at all levels. The process of government informatisation is not merely an upgrade at the technical level; it also represents a profound change in the governance model [4]. The utilisation of information technology enables the government to integrate and disseminate resources in a more efficacious manner, facilitating the precise alignment of services with individualised requirements and thereby enhancing the quality and efficiency of public services, while concomitantly augmenting the transparency and social credibility of the government. Furthermore, the establishment of an integrated government service platform will facilitate the elimination of information silos, optimise government processes, reduce service costs and enable the public to access a range of government services in a more convenient manner [5].

It can be argued that the promotion of government informatisation and integrated platforms is not only a necessary measure to respond to the trend of digital transformation, but also a key step to improve government governance capabilities, enhance public service levels and promote social equity and sustainable development. In this process, it is crucial to consider the varying needs of different groups while maintaining efficiency, particularly with regard to the accessibility of groups with low literacy rates and special needs, which has emerged as a significant challenge in the construction of modern government affairs platforms [6].

As a crucial conduit for disseminating official information and providing government services to the public, the integrated government platform is an indispensable component of the advancement of e-government. The platform plays a pivotal role in enhancing the efficiency of government operations, reducing operating costs, and optimising office processes [7]. Furthermore, it has a profound impact on the improvement of public services, the enhancement of the government's image, the promotion of government reform, and the strengthening of public supervision. The construction of an integrated government affairs platform will facilitate the acceleration of the construction of a service-oriented government, the promotion of the process of government information management, and the modernisation of the social governance system and governance capacity [8].

The study of the integrated government platform facilitates the optimisation of service processes and the innovation of service modes. Furthermore, it provides theoretical support and technical pathways for the construction of a more efficient, convenient and transparent service-oriented government. The platform can assist the government in addressing the current issues of "difficult, slow, and complicated," as highlighted by the public, and facilitate the transformation of government services into a more intelligent model. By further integrating the Internet with government services, the platform can facilitate real-time updates and accurate notifications, aligning government services with the needs of the people [9].

2. Related Work

In their 2020 study, Jithesh, Anupriya, and Satish [10] examine the intricate interrelationship between the proliferation of virtual social networks (VSNs), e-government maturity, government administrative effectiveness, and corruption from a multitude of theoretical perspectives. Their analysis draws upon the value framework of e-government impact assessment, integrating insights from technological determinism, general deterrence theory, and Habermas' public sphere perspectives. They put forth a comprehensive conceptual framework that posits that enhanced e-government maturity can be an efficacious means of reducing corruption within the legislative, executive, and judicial branches.

The study demonstrates that as e-government systems mature, there is a notable increase in transparency, accountability and public participation in government. This not only enhances the efficiency of government administration but also plays a role in curbing corruption. The extensive utilisation of virtual social networks enables the government to more effectively disseminate information, enhance policy transparency and facilitate social supervision, thus propelling the modernisation of the entire social governance system. This study offers a comprehensive examination of the role of e-government in anti-corruption and the enhancement of government governance capabilities. It provides a theoretical foundation for governments to promote the development of e-government and improve administrative efficiency.

In their 2022 article, Grinin L and Korotayev [11] posit that the ongoing global pandemic has served as a significant catalyst for the accelerated transition of e-government towards the formation of an "e-state". In the context of a global public health crisis, the pandemic has compelled governments to rapidly adopt digital solutions to guarantee the continuity and efficacy of government services. In particular, during the initial stages of the pandemic, when social distancing and lockdown measures were in place, e-government became the sole means of ensuring the continuity of government services, thereby accelerating the digital transformation of governments.

3. Methodologies

A. Multimodal Learning

In order for the interface to be adaptable to the user's specific requirements, it is essential that the system is able to comprehend the user's state and needs through multimodal learning. Multimodal learning is a method of combining different types of information inputs, including visual, speech, and text, in order to provide a more comprehensive analysis of the user.

Let us consider a scenario in which the user's input data comprises visual information (denoted as V), speech information (denoted as S), and text information (denoted as T). Each of these can be represented as eigenvectors by a deep learning model. The aforementioned eigenvectors can be defined as follows, as illustrated in Equation 1, Equation 2 and Equation 3:

$$V = f_{vis}(I_{img}), \tag{1}$$

$$S = f_{aud}(I_{audio}), \tag{2}$$

$$T = f_{txt}(I_{txt}), \tag{3}$$

The I_{img}, I_{audio}o, and I_{txt} represent the input image, audio, and text data, respectively. The f_{vis}, f_{aud}, and f_{txt} are deep neural networks that extract visual, audio, and text features, respectively. After the multimodal feature fusion, the comprehensive feature vector X_{user} can be obtained by weighted average, as demonstrated in Equation 4:

$$X_{user} = \alpha V + \beta S + \gamma T, \tag{4}$$

The values of α, β, and γ are automatically learned through the training process. The fusion process enables the model to obtain comprehensive user status information and conduct a detailed analysis of the user's personalised requirements. Figure 1 shows the specific framework of our proposed model.

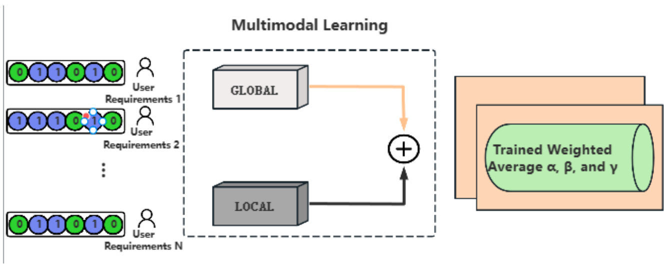

Figure 1. Illustration of Multimodal Learning.

B. Reinforcement Learning

In order to achieve dynamic interface adjustment, this study employed the use of reinforcement learning (RL) mechanisms. In particular, the system employs reinforcement learning to optimise the interface layout, text size, colour contrast and voice assistance features, thereby ensuring that each user has the optimal experience when interacting with the system.

The environment state is defined as s_t, which indicates the user's demand state at time step t. The action space, designated as Action Space A_t, encompasses a range of options, including interface layout adjustment, font size change, colour contrast adjustment, and voice assistance enablement. The action at that moment, designated as Action $a_t \in A_t$, is indicated by the selected action by the system. The objective of the system is to optimise the reward function R_t, which quantifies the impact of interface alterations based on user feedback.

The Q-learning algorithm of reinforcement learning is employed to define the Q-value function $Q(s_t, a_t)$, which represents the expected cumulative reward obtained after taking action a in state s_t, as illustrated in Equation 5:

$$Q(s_t, a_t) = R_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}), \tag{5}$$

In this context, R_t represents the immediate reward, whereas γ denotes the discount factor, which indicates the relative importance of the future reward. Reinforcement learning algorithms are designed to identify the optimal action by continually updating the Q-value function, thereby enabling the automatic adaptation of the interface. The specific update formula is given by Equation 6:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(R_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t), \tag{6}$$

Among the parameters, α represents the learning rate, which determines the extent to which the newly introduced information affects the Q value. By means of continuous iteration and updating, the system is capable of automatically selecting the most appropriate interface adjustment scheme in accordance with the varying user demand states, thereby achieving a highly personalised user experience.

C. Adaptive Interface Adjustments

In order to gain a more accurate understanding of the user's needs and make real-time interface adjustments, this study integrates sentiment analysis and speech recognition modules. The sentiment analysis module employs a deep neural network, such as BERT or LSTM, to analyse the user's text input or speech sentiment state, thereby assisting the system in determining the user's sentiment and needs. To illustrate, in the event that a user is experiencing distress during their interaction with the system, the system can alleviate the user's emotional burden by increasing the frequency of voice prompts or modifying the layout of the interface. The process of sentiment analysis can be represented by the following Equation 7:

$$\hat{y} = f_{sentiment}(T), \tag{7}$$

In this context, T represents the user's text input, whereas $f_{sentiment}$ denotes the sentiment analysis model. The predicted sentiment label, represented by $\hat{y}$, may assume one of three possible values: "positive", "neutral" or "negative".

The speech recognition module serves to enhance the real-time and accuracy of interface adjustment by converting the user's voice input into text. In the event that the user's speech input is I_{audio}, the speech recognition model is capable of converting it to text via the following Equation 8.

$$T = f_{sspeech}(I_{audio}). \tag{8}$$

4. Experiments

A. Experimental Setups

The data sources employed in this study encompass a range of data types, including simulated user behaviour data, data pertaining to users with low literacy rates and special needs, generated virtual user needs data, and existing public datasets. In particular, virtual behavioural data can be generated in a number of ways. Firstly, a simulated user interaction model can be constructed. Secondly, experimental scenarios can be designed to simulate the operational behaviour of users with

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 December 2024 doi:10.20944/preprints202412.1062.v1

low literacy rates and special needs. Thirdly, natural language processing technology can be employed to generate user demand data. Finally, experimental data can be supplemented with existing public datasets and user feedback data. Following Figure 2 shows the proposed simulation platform.

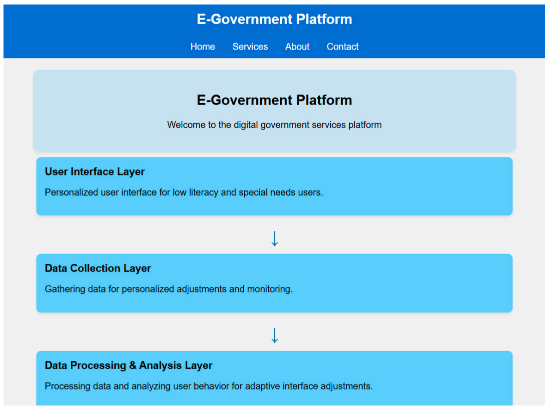

Figure 2. Illustration of Proposed Platform.

B. Experimental Analysis

The primary evaluation metric employed is accuracy, which is used to assess the model's ability to accurately identify and predict user needs under varying user numbers and to make corresponding interface adjustments. Figure 3 presents a comparative analysis of the performance of the proposed method with multiple established models, including CNN, LSTM, Transformer, and a baseline, to assess the stability and efficiency of the model under varying loading conditions.

As illustrated in Figure 3, the accuracy of the proposed method has remained consistently high across all user counts, exhibiting a slight decline but maintaining stability, thereby substantiating its superiority and adaptability in high-user-load scenarios. In contrast, the performance of the benchmark model is markedly different. The accuracy of the CNN and LSTM models decreases significantly with the increase of the number of users, indicating that their efficiency and accuracy in the context of large-scale user requests is compromised. The performance of the Transformer model is comparable to that of the proposed method, yet a discernible discrepancy remains. Baseline exhibits lowest accuracy with pronounced decline as number of users rises, indicating that the model is particularly susceptible to performance degradation under high loads.

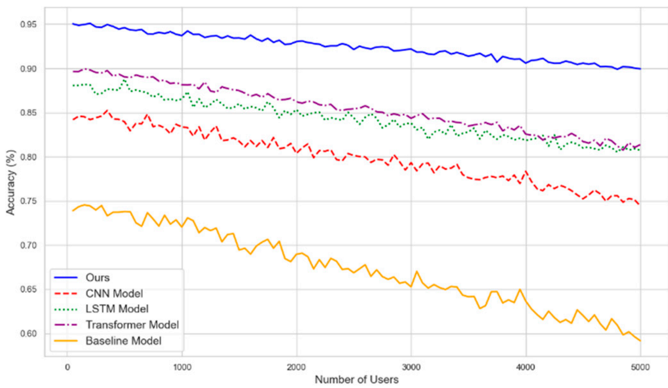

Figure 3. Accuracy Comparison of Ours vs Benchmark Models.

As an evaluation index, interface adaptability is employed to ascertain whether the system is capable of dynamic interface adjustment in accordance with the user's requirements, with the objective of enhancing the interactive experience.

Figure 4 depicts the distribution of fitness scores for the five methods. The data distribution for each method is presented in the form of a box plot, with a median line representing the median fitness score for the method. The upper and lower bars represent the maximum and minimum values of the data, respectively, while the median line represents the median fitness score for the method. As illustrated in the Figure 4, the Ours method exhibits a higher and more stable adaptability score, demonstrating superior performance compared to other methods, particularly of range and median performance for upper and lower quartiles, which indicate enhanced interface adaptability.

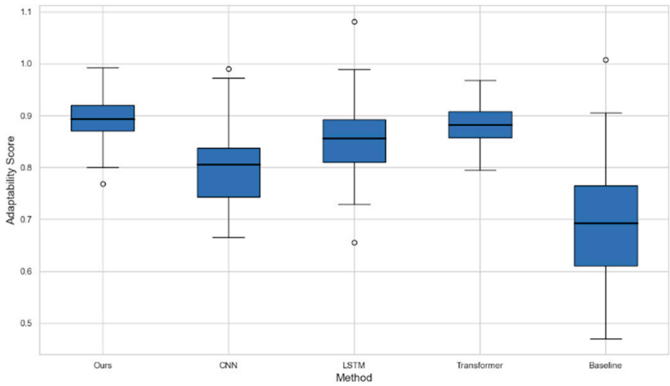

Figure 4. Interface Adaptability Comparison.

In Table 1, we compare the computational performance of the four algorithms. Ours is our proposed custom model, which has been optimized to exhibit high accuracy and low training time, and is suitable for tasks that require efficient computation and accurate prediction. In contrast, CNN, as a classical convolutional neural network, has a significant advantage in processing image and spatial features, although the accuracy is slightly lower than that of Ours, and the training time and FLOPs are higher. LSTMs perform well when working with time series data and can capture long-term dependencies, but their training time and memory usage are large, and FLOPs are relatively high.

Table 1. Computation Metrics Comparison Results.

Algorithm	Accuracy (%)	Training Time	FLOPs (Billion)	Memory Usage (MB)
Ours	92.5	120	2.5	800
CNN	89.3	150	3	950
LSTM	85.7	200	4.2	1100
TRANSFORMER	88.2	100	1.8	750

5. Conclusions

In conclusion, we propose an adaptive interface model based on deep learning technology, with the objective of enhancing the accessibility and user experience of e-government platforms, particularly for users with low literacy rates and special needs. Experimental analysis serves to corroborate the stability and adaptability of the proposed approach in different scenarios. Results demonstrate that integration of deep learning and multimodal mechanisms can significantly enhance the accessibility e-government platforms, establishing a foundation for a more inclusive and user-friendly public service system.

 doi:10.20944/preprints202412.1062.v1

References

1. Moulin, Claude, and Marco Luca Sbodio. "Improving the accessibility and efficiency of e-Government processes." 9th IEEE International Conference on Cognitive Informatics (ICCI'10). IEEE, 2010.
2. Bonacin, Rodrigo, et al. "Accessibility and interoperability in e-government systems: outlining an inclusive development process." Universal Access in the Information Society 9 (2010): 17-33.
3. Abdurahman, Saleem. "Exploring the challenges and implications of designing, developing and implementing accessible e-government services." (2022).
4. Ramadoan, Syahri, et al. "Optimizing public accessibility: A review of the urgency of e-government for Bima District." The International Journal of Politics and Sociology Research 12.1 (2024): 1-9.
5. Qi, Min, and Junshu Wang. "Using the Internet of Things E-Government Platform to Optimize the Administrative Management Mode." Wireless Communications and Mobile Computing 2021.1 (2021): 2224957.
6. Desmal, Abdulla Jaafar, et al. "A user satisfaction model for mobile government services: a literature review." PeerJ Computer Science 8 (2022): e1074.
7. Halsbenning, Sebastian, et al. "Playing (government) seriously: design principles for e-government simulation game platforms." Innovation Through Information Systems: Volume III: A Collection of Latest Research on Management Issues. Springer International Publishing, 2021.
8. Ferraz, Reinaldo, et al. "A platform to check website compliance with web accessibility standards." Proceedings of the 20th International Web for All Conference. 2023.
9. Uutsi, Teopolina, and Tulimevava K. Mufeti. "An investigation of e-government web accessibility for visually impaired persons in Namibia." Resilience, Entrepreneurship and ICT: Latest Research from Germany, South Africa, Mozambique and Namibia (2021): 403-420.
10. Arayankalam, Jithesh, Anupriya Khan, and Satish Krishnan. "How to deal with corruption? Examining the roles of e-government maturity, government administrative effectiveness, and virtual social networks diffusion." International Journal of Information Management 58 (2021): 102203.
11. Grinin, Leonid, Anton Grinin, and Andrey Korotayev. "COVID-19 pandemic as a trigger for the acceleration of the cybernetic revolution, transition from e-government to e-state, and change in social relations." Technological Forecasting and Social Change 175 (2022): 1213