International Journal of Scientific Research in Science, Engineering and Technology

Print ISSN - 2395-1990
Online ISSN : 2394-4099

OPEN ACCESS

Available Online at : www.ijsrset.com
doi : https://doi.org/10.32628/IJSRSET2411117

The Use of Voice Recognition AI to Make It Easier for Elderly People to Use Mobile Phones

Dr. Rajendra Kumar Mahto

Assistant Professor, Department of Information Technology, Dr. Shayama Prasad Mukherjee University, Ranchi, India

ARTICLEINFO

Article History :

Accepted: 02 Feb 2024
Published: 14 Feb 2024

Publication Issue :
Volume 11, Issue 1
January-February-2024
Page Number :
209-213

ABSTRACT

For a considerable amount of time, people have been aware of the issue that older people find it difficult to operate mobile phones and frequently decline to connect with modern technology. The social phenomenon of elder rejection of smart devices, including smartphones, has already been established. The main cause is that the majority of smartphones have complicated systems that are too hard for elderly people to understand. The elderly are desperate to get rid of the barrier preventing their requests from being entered into the system. Conversely, artificial intelligence speech recognition can essentially disregard this problem because it can replace the manual manipulation portion with command comprehension based on machine learning. AI speech recognition can analyze voice commands, identify key words to form into an order, and ultimately carry out the order correctly thanks to the deep learning techniques covered in this paper. Voice recognition software can automatically react quickly and identify the health conditions of the elderly based on characteristics of their voice. The goal of this article is to discover how to create AI voice recognition systems that are specifically tailored for a particular population, such as senior citizens, who are the system's intended audience. Additionally, this article offers logical predictions for the use of AI voice recognition in the fields of psychological research and mental health treatment.

Keywords : Voice Recognition, Speech Recognition, Artificial Intelligence, Phone.

I. INTRODUCTION

Following the development of smartphones, the sometimes confusing and intricate mobile phone systems have caused problems for a lot of smart phone users. Certain systems have a large number of options, which complicates ordering. Certain systems make it difficult to read sentences because they have to use smaller words to fit the large number of words on the screen; other systems may have unclear system descriptions that prevent phone users from placing precise orders. As our society ages, it is imperative that we enhance the phone-using experience for seniors and ensure that they can use it without difficulty. The issue of some elderly phone users finding mobile phone usage too complex has long been recognized by mobile phone companies. The elderly have given up attempting to use smartphones regularly in their everyday lives as a result of the challenges or annoyances associated with using their phones. Phone-using improvements and simplifications have long been a problem for companies and system-developing groups. Larger wording and AI voice assistants are two practical ways to address this issue. By reducing the discrepancy between the elders' requests and the computer orders' inputs, they have enhanced the elders' phone-using experience by enhancing the system's comprehension of their requests. These days, voice assistants are already highly "intelligent." They can function well for the majority of adults or children and can translate voice commands into the appropriate word orders. These days, it's reasonable to argue that AI voice assistants have simplified phone use beyond recognition.

However, current AI voice assistants are too general for the elders, since general voice assistants frequently misunderstand the voice commands given by the elders. According to a recent survey, the general voice assistant misunderstands the majority of commands given by the elderly when they activate apps and translate voice to text. This is due to the fact that elderly people have different needs than those of adults and children, and they also have different voice tones, slower speech rates, speaking habits, and even intonations. This would lead to misunderstandings of the words of the elderly, which would greatly reduce the effectiveness of voice assistants. For example, they examine fake noises because they lack the lower tones of elderly people and instead have general tones. Because elderly people speak at a slower rate than younger people, they might interrupt during a speech and mispronounce words that older people might pronounce differently. than the general public, etc. Older smartphone users have been experiencing these issues, which has led to AI voice assistants being far less useful than they should be.

The primary goal of this research is to determine how to effectively interpret elders' orders into a voice recognition system. Additionally, by allowing a system to adapt new training methods, it becomes possible to develop a voice recognition system that is specifically intended for elders.

A potential fix is offered in the article. When developing the training system, it is to make use of specific training data, targeted training methods, and various test standards, all of which will be discussed in the following sections.

II. LETERATURE REVIEW

Edward Porter, an inventor, created the first practical voice recognition system in 1983. Every single the construction of subsequent voice recognition systems was similar to that of this one: recording verbal inputs, converting verbal strings into word strings, extracting keywords from word strings, and using the keywords to construct sentences.[1]

Hisayuki Nagashima, a researcher, presented a particular technique for extracting keywords from spoken inputs in 2008. He provided a clear analysis of the tools and techniques that will be applied to the voice recognition procedure. [2]

Researcher Sean Doyle described how to allow a voice recognition system to make improvements on its own in 2001. This can be considered as the first essential step in the development of artificial intelligence-based voice recognition systems. This has offered guidance for all speaker-based or combination of speaker and requires ongoing enhancements. This is essential to the development of the system that this article discusses.[3]

A study conducted in 2011 demonstrated how much the user's language proficiency affected the performance of a particular AI voice recognition system. This has brought up a brand-new subject: how can systems be developed for users with limited language skills? Meanwhile, the goal of this article is to figure out the system's solution.[4]

The system and approach that can be applied to a voice recognition system has been determined by researcher Paul Angott. He noted that systems can be programmed to play back pre-recorded audio patterns in his theorem. These patterns would rank highly in the voice recognition algorithm, which would encourage attempts to confirm these patterns. [5]

III. ANALYSIS

To develop a voice recognition system specifically for the elderly, research into the lifestyle patterns of the elderly is required. It has been discovered that social interactions and activities play a significant role in how elders respond to their phones. They also have a higher need for health-related apps than do regular phone users, and they don't need as many entertaining apps like game and video websites. In addition, their speech patterns are more varied and they speak at lower rates and tones. As previously mentioned, all of these characteristics resulted in an entirely different and much larger database than typical voice recognition training systems, which led to unique training data, focused training techniques, and distinct test standards.

3.1. Instructional material

The majority of the training data is made up of Six sections: works, entertainment, emergency calls, health-related instructions, word contacts, phone calls, and work. The percentage of word contact section orders is significantly higher than that of other sections, per the survey. Thus, additional test data are required to give this section additional accuracy. The amount of test data is decreased for the other four sections—the phone call section, the health-related demand section, the work app section, and the entertaining app section—because their calling frequencies are decreasing. On the other hand, unlike other sections, the frequency of calling out the emergency section could not be generally measured. For this section, additional data are also supplied because inaccuracy could have very negative effects. In the meantime, the weights assigned to the six sections would vary based on both the major information and later understanding of the user. The weights vary based on the user's usage habits, such as how frequently a specific type of section is mentioned or how frequently the voice command is entered again into the application.

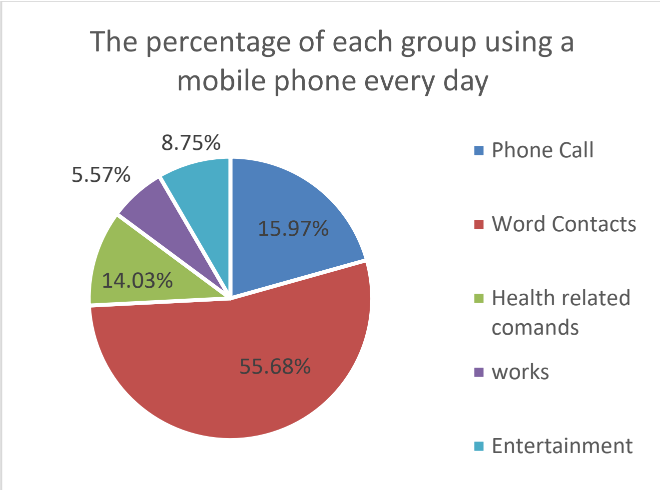

Figure 1 : Daily mobile phone usage percentages for each section

Figure 1 was produced using data from a survey of 1000 senior citizens in Shanghai that was carried out in neighbourhoods and nursery homes. It has displayed the estimated frequency at which each section's key words are expected to be used. Meanwhile, this has established a number of fundamental coefficients while constructing establishing the model, such as the quantity of database required to train each keyword section, the section weights, the accuracy requirements for each section to deliver optimal user experience, etc.

3.2. Instructional strategies

Converting voice commands into computer orders would involve two main steps:

The target sound track must be separated from background noise in the first step. To select the needed sound track for this section, the system would first split the target data into multiple sound tracks. It would then use K-Nearest Neighbour methods, which take into account attributes like volume, sound clarity, tones, and speech rates. The K-Nearest Neighbour method is just as accurate when compared to other modules like vector learning and so forth. The coefficients are much simpler. Adapt to change, and the K-Nearest Neighbour approach would give the system greater flexibility. For instance, the K-Nearest Neighbour method allows the system to make changes in the smallest number of attempts, such as increasing the working section's weight by only adjusting the weight coefficients, if the user calls out working section orders more frequently.

The second step involves selecting important terms from the audio track in order to create a computerized ranking. In this section, the system would first segment the audio file into words, and after that, it would identify key words by counting the likelihood that each word that was separated would be a key word using Bayesian learning. It is important to remember that this learning process could lead to various outcomes based on how much weight each section of the data has. Since it provides clear reflections of probabilities and may be able to provide multiple ordering options based on the final probabilities of each word being a key word, Bayesian learning has already been proven to be the most effective technique. For example, the system would provide three different orders if it detected two potential key words in the same voice order. The user could then select to activate the first order, the second order, or both orders.

The process of valuing speech rates exists because elderly people typically speak more slowly when speaking, which makes this training method different from training methods for general voice systems. instructions to the voice recognition AI system. As a result, the two systems' definitions of "generality" may differ, with the former allowing for pauses or slower speech rates while the latter does not.

3.3. Standards for testing

The test standards for this system are different from those for general voice recognition systems in that they are elder-based and demand a higher degree of fitting because the system may be able to save lives in emergency situations. Over 95% accuracy is the anticipated rate.

3.4. Test outcomes

The system's testing and training did not result in a working voice recognition system because of the limitations of computing power. However, a few basic studies came to the following conclusion:

During the process of moving from verbal inputs to word strings, five is the primary recognition frequency required to accurately capture the intended verbal input. To decide which audio track should be examined, those with frequencies close to the primary recognition frequency will be given priority.

The ideal range for the primary recognition frequency is 350–500 Hz. This is due to the fact that most elderly people speak at a frequency of 350–400 Hz. The accuracy peak that is the highest is situated between 370 and 380 Hz. The second-highest peak is found between 430 and 450 Hz. This is most likely due to the fact that the majority of male and female participants speak at frequencies that are close to the first and second peaks, respectively.

IV. CONCLUSION

In the future, it may be possible to turn the voice-transferring technique for elderly people into an app, which would improve and simplify their phone-using experience and help them gradually come to terms with smartphones. Techniques for voice recognition software or applications targeted at particular demographics have enormous potential in the future. Comparable techniques can also be applied to help children or individuals with mental illnesses like autism use their phones more effectively. It would probably help us learn more about the worlds inside their minds in addition to being a more convenient way for them to use their phones.

It is possible to combine the emergency and health-related commands sections into a single section that primarily acts as a health condition indicator for users. The recently added "health section" can to identify possible illnesses or health risks, gather data on the user's frequency of pausing, the volume of the trembling sound, or how the voice sounds different on typical days. When combined with additional thermal modules and pulse valuation capabilities, the voice recognition module may even be able to perform some of the standard health checks, sparing the elderly the trouble of traveling to the hospital for a checkup. Meanwhile, the prior emergency section would continue to be responsible for automatically initiating emergency calls and displaying first aid instructions on the phone's screen.

V. REFERENCES

[1]. Porter, Edward W. Boston, Mass. voice recognition software. US Patent. US4829576A. Publication number: May 9, 1989. Patent: US4829576A; https://patents.google.com/patent/en#title

[2]. Wako and Hisayuki Nagashima. voice identification software, voice recognition hardware, and voice recognition technique. US Patent. Publication ID: US8005673B2. Date of publication: August 23, 2011. Patent: US8005673B2/en?oq=US8005673B2 can be found at https://patents.google.com.

[3]. Sean Doyle. enhancing a voice recognition system automatically. US Patent. US7103542B2. Published September 5, 2006. url=US7103542B2/en at patents.google.com/patent/US7103542B2

[4]. John D. E. Gabrieli, Stephanie N. Del Tufo, and Tyler K. Perrachione. Science. [J] Vol.333, Issue 6042, pp. 595, July 29, 2011 10.1126/science.1207327 is the DOI.

[5]. Angott, Paul. Technique and apparatus for voice recognition. Publication number: 13 August 2013 US8510103B2. The patent number is US8510103B2/en at patents.google.com.

Cite this article as :

Dr. Rajendra Kumar Mahto, "The Use of Voice Recognition AI to Make It Easier for Elderly People to Use Mobile Phones", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 11 Issue 1, pp. 209-213, January-February 2024. Available at doi : https://doi.org/10.32628/IJSRSET2411117

Journal URL : https://ijsrset.com/IJSRSET2411117