South Asian Journal of Engineering and Technology Research Article

Open Access Full Text Article 121(2022) 74-83 DOI: 10.26524/sajet.2022.12.012

Integrating of Voice Recognition Email Application System for Visually Impaired Person using Linear Regression Algorithm

Glenn Arwin M. Bristol[a], Joevilzon C. Calderon[b]

[a] **Taguig City University- College of Information Communication and Technology Department**

[b] **Taguig City, Philippines**

Corresponding author.
Correspondence: Glenn Arwin M. Bristol
Email: glennarwinmacalinaobristol@gmail,com

Article info
Received 26th December 2021
Received in revised form 20 February 2022 Accepted 22 March 2022

Keywords
Visually Inpaired, Voice Recognition, Integration

Abstract
The outcome of this study will surely help visually impaired people, who face difficulties in accessing the computer system. Voice recognition will help them to access e-mail. This study also reduces cognitive load taken by a visually impaired users to remember and type characters using keyboard. If this system is implemented, self-esteem and social and emotional well-being of the visually impaired users will be lifted up for they will now feel they are being valued in the society and has fair treatment and access in technologyThe main function of this study is to use a keyboard of the user that will respond through voice. The purpose of this study is to help a visually impaired person to use modernize application to interact with voice recognition system with the use of email into different types of modern gadgets Line computers or mobile phones. In terms of Functionality of the application, the proponents will use a set of APIs' or Application Program Interface such as Google Speech-to-text and text-to-speech application and it will process though Email System and also the SNMTP or Simple Network Management Protocol will be used for mailing services, in programming software, the proponent will be using PHP for the backend of web interface. HTML and CSS is the front end programming for the creation of Web Base User Interface that will be used. Voice typing and Dictation Speech Interaction models using windows dictation engine. The proponent used descriptive research design in this study. Descriptive research design is being used by the proponents to describe the characteristics of a population or phenomenon of visually impaired persons being studied. Descriptive research is mainly done because the researchers wants to gain a better understanding for a topic. It focuses on providing information that is useful in the development. The research is based on mixed method focused in producing such informative outcomes that can be used. Based on the results of the surveys, conclusions were drawn: Majority of the respondents were male adultery period ranging ages 32-41.all are working in the massage therapist. Majority of the respondents rated the overall function of the application Excellent and rated the level of security of the application is Secured.

I. INTRODUCTION

Nowadays, not all people are aware on how a typical Blind person suffers in their everyday living. According to (Global Data on Visual impairments 2012). Many of them are always problematic on how they live with difficulties with normal daily activities such as driving, reading, socializing and walking. And to boost their confidence for interacting and accessing different technologies. To define what is Visual impairment it describes a person who has a vision loss and someone who cannot see at all or has a partial vision loss, completely blind, and others called legal blindness. Based on the World Health Organization Prevention of Blindness and Deafness 2008 edition. Globally there are 2.2 billion people having a near or distance vision impairment and the estimated numbers of visually inpaired persons in 2012 weere 286

million worldwide of which 246 million had low vision and 39 million were blind. Of those who are blind 90%. According to (Althomali T (2012) The main cause of blindness and vision impairment are they have cataracts uncorrected refractive errors in their eyes and caused by things like infection, genetics or injury. The vision of visually infaired and blind measures 20/200 or worse However, the majority of people ages 50 years have a vision impairment and blindness so vision loss can affect people of all ages.

Moreover, Internet has made life of people so easy that people today have access to any information they want while sitting at their home. One of the main fields that Internet has revolutionized is "communication". While talking about communication over Internet, the first thing that comes in our mind is E-mail. E-mails are considered to be the most reliable way of communication over Internet, for sending or receiving some important information.

Voice Recognition can be of great value to blind people as it empowers them to independently complete tasks which they would normally not be able to accomplish, boost their self-confidence and Visually impaired people can do a lot of things such as reading mails or managing their own accounts without help from others with the use of voice recognition, Voice recognition serves as an alternative to typing on a keyboard. It gives users the option to use a speech-to-text system, which means users talk to the computer and their input is interpreted and converted to electronic text. Put simply, you talk to the computer and your words appear on the screen. The proponent believes that the increasing number of blind persons in the Philippines will begin building a demand for the new trends in technology. It seems that the time is right to begin developing high-technology services to meet the needs of this growing population of literate blind not only for them to consider themselves in to the millennials but also a big chance to boost their social status and self-esteem. This study is in the hopes of better serving this hidden population of blind people. We hope to gain new ideas and make new innovations that will benefit all those who are blind in the Philippines.

II. LITERATURE REVIEW

Based on a coduct Survey article on the Use of Mobile Applications for People who Are Visually Impaired (Griffin, Banda, Ajuwon 2017), The few studies indicated and conducted to the person with visual impairments or who are blind and having a low vsion are concerning in the mobile application usage. The use of mobile application become increase in globallycompleted the online survey the summary with the total of 259 participants having visual impairments. They use Descriptive statistics and bivariate tests methods where they used to examine all associations for the demographic's characteristics and mobile app use. The result of the online survey is that all the partcipants are rated special apps with useful (95.4%) and accessible (91.1%) tools for individuals with visual impairments. In middle age adult group More than 90% are strongly agreed with the practicality of special apps while in old adult groups are significantly higher percentage was observed. Some of the participants with a low vision as addition to the result considered special appl has a lesss accessible for using apps. For the final results of this study this population is satistisfied with the us of mobile apps and they would like to see the improvements, furthermore, the result show that the persons who have visual impairments frequently use application design to thm to accomplish different daily activities. For the developing of the application for visual impaired and blind is use to refine and tests the existing apps, so the study provides the preliminary information regarding the app's usage for the visually inpaired and blind.

Morover, based on the study, Ecommerce Based Online Shopping for Visually Impaired People using Speech Recognition*)* on pointing out the use of online shopping for visually impaired, For Visual impaired people it is difficult to do online shopping (Kunal, Nawkhare 2017*)* We are developing system which helps blind people to choose color of clothes along with categories such as material, size, patterns etc. by using Automatic speech recognition module. Translating of spoken words into text using deep learning method". According on their study, it successfully replaced Gaussian mixture speech recognition and feature coding at an increasingly larger scale. The main aim of this project is to propose an online shopping application especially for blind people through which they can choose whatever they want by simply speaking a sentence and to implement this we are going to integrate deep learning method and clustering.

On Other study in related to the Speech- Voice recognition *(Gros, Mihelic, Pavesic 2017)* "HOMER is a voice-driven text-to-speech system developed for blind or visually impaired persons for reading the Slovenian texts". Based on their study, users can obtain texts from the special corpora organised on the computer network server at the information center of the Association of the Slovenian Blind and Visually Impaired Persons. The system consists of three main modules. The text-to-speech module enables speech synthesis from an arbitrary Slovenian text input, the speech recognition module performs speaker independent isolated word recognition and the dialogue module controls the different tasks of the HOMER system and obtains texts from the source text corpora. Presently, the system runs under Linux and requires a Pentinum/133 PC with minimum 32 MB of RAM and an additional standard 16-bit sound card. While according to M. Bazzani and E.M. Mumolo, (Alcatel FACE Res. Center, Pomezia, Italy), they stated: "PC-based telephone communications system for deaf-blind people is a system has been developed that allows deaf-blind people to communicate with others by a standard telephone set". Based on their study, the system will use speech processing technologies integrated in a PC (personal computer). The handicapped person can

activate or receive a telephone communication, entering sentences by typing on a Braille terminal and receiving the messages on a Braille tactile display. A text-to-speech subsystem converts the typed sentences into voice, and a connected-word voice recognizer converts the vocal answer into text ande then into a Braille message. Results of a field trial of the system carried out in an Italian community of handicapped people are reported.

III. RESEARCH METHODOLOGY

The researcher used descriptive research design in the project. Descriptive research design is being used by the proponents to describe the characteristics of a population or phenomenon of visually impaired persons being studied. Descriptive research is mainly done because the researchers wants to gain a better understanding for a topic. It focuses on providing information that is useful in the development. The research is based on mixed method focused in producing such informative outcomes that can be used, Respondents of the study

The respondents of this study were the visually impaired persons who can able to speak and hear to be able to verify all the important information on his/her email account and use the voice recognition application properly.

The proponent used one method for data gatehering is the interview as the core and umbrella of this study. Interviews on the respondents which are some of the visually impaired persons who uses desktop computer, a visually impaired person who is engaged in office works who has problems in sending mails, a visually impaired person who also has experiences in using voice recognition AI like Siri. For computing the percentage of the respondents, Statistical Tools and Instruments are being used for determining the sampling Population, Frequency, Mean, Standard Deviation and Annova Analysis for the Validation and distribution of the instruments used.

IV. FINDING AND DISCUSSION

Table 1 – The Demographic Profile of the Respondents in Terms of Age Bracket

Age Bracket	Frequency (f)	Percentage (%)	Rank
18 – 25	0	0%	5
26 – 31	11	22%	3
32 – 41	21	42%	1
42 – 49	13	26%	2
50 above	5	10%	4
Total No. Respondents (N)	50	100%	

As shown from the table, out of 50 respondents in terms of age bracket, 32-41 got the highest frequency of 21 or 42 % is rank 1, 42-49 with a frequency of 13 or 26 % is rank 2, 26-31 with a frequency of 11 or 22 % is in rank 3 while 50 above with a frequency of 5 or 10 % is in rank 4 and 18-25 got no frequency of 0 or 0 % is in rank 5.

Table 2 – The Demographic Profile of the Respondents in Terms of Gender

Gender	Frequency (f)	Percentage (%)	Rank
Male	34	68%	1
Female	16	32%	2
Total No. Respondents (N)	50	100%	

As shown from the table, out of 50 respondents in terms of gender, the frequency of Male is 34 which is equivalent to 68% which is rank 1, while the frequency of the Female is 16 or 32% which is in rank 2.

Table 3 – The Demographic Profile of the Respondents in Terms of Job/Designation

Job/Designation	Frequency (f)	Percentage (%)	Rank
Singer	12	24	2
Massage therapist	16	32	1
Student	10	20	3
Radio announcer	**2**	**4**	**5**
Factory worker	**2**	**4**	**5**
Call center agent	**8**	**16**	**4**
Total No. Respondents (N)	**50**	**100%**	

As shown from the table, out of 50 respondents in terms of Job description, the frequency of Massage therapist got the highest of 16 or 32 % which is rank 1, Singer with a frequency of 12 or 24 % which is rank 2, Students with a frequency of 10 or 20% which is in rank 3 while Call center agent with a frequency of 8 or 16% which is in rank 4 and both Radio announcer and Factory worker got the frequency of 2 or 4% which is in rank 5.

Table 4 – The Demographic Profile of the Respondents in Terms of Highest Educational Attainment

HIGHEST EDUCATIONAL ATTAINMENT	Frequency (f)	Percentage (%)	Rank
College Level	**3**	**6%**	**3**
Vocational	**32**	**64%**	**1**
High School Graduate	**0**	**0%**	**4**
High School Level	**0**	**0%**	**4**
Elementary Graduate	**0**	**0%**	**4**
Elementary Level	**15**	**30%**	**2**
Total No. Respondents (N)	50	100%	

As shown from the table, out of 50 respondents in terms of Highest educational attainment, the frequency of Vocational got the highest of 32 or 64 % which is rank 1, Elementary level with a frequency of 15 or 30 % which is rank 2, College level with a frequency of 3 or 6 % which is in rank 3 while High school graduate, high school level, and elementary graduate got no frequency.

Table 5 – The Demographic Profile of the Respondents in Terms of Experience using any kind of gadgets

Experience using any kind of gadget	Frequency (f)	Percentage (%)	Rank
1 to 5 years	**10**	**20%**	**2**
6 to 10 years	**30**	**60%**	**1**
Not at all	**10**	**20%**	**2**
Total No. Respondents (N)	50	100%	

As shown from the table, out of 50 respondents in terms of Experience using any kind of gadgets, the frequency of 6 to 10 years got the highest of 30 or 60 % which is rank 1, 1 to 5 years got a frequency of 10 or

20 % which is rank 2, while those who do not have experience using any kind of gadget got a frequency of 10 or 20 % which is also rank 2.

Part II. Evaluation of the Respondents

Table 1 – The evaluation of the respondents in terms of Accessibility no. 1

Criteria Accessibility	**How would you rate the Log-in method in terms of bounding g-mail accounts?**		
Likert Scale	**Frequency f(x)**	**Percentage (p)**	**Rank**
Excellent	21	42%	1
Very Good	14	28%	3
Good	15	30%	2
Fair	0	0%	4
Poor	0	0%	4
Total no. Respondents (N)	**50**	**100%**	
Mean	**4.12**		
Verbal Interpretation	**Very Good**		

As shown from the table, out of 50 respondents in verbal interpretation in terms of accessibility. Excellent got the highest frequency of 21 or 42% which is in rank 1, Good with a frequency of 15 or 30% which is in rank 2, Very good with a frequency of 14 or 28% which is in rank 3 while Fair and Poor got no verbal interpretation which is rank 4.

Table 2 – The evaluation of the respondents in terms of Accessibility no. 2

Criteria Accessibility	**How would you rate the accessibility of the application in terms of creating and sending messages through voice recognition?**		
Likert Scale	**Frequency f(x)**	**Percentage (p)**	**Rank**
Excellent	21	42%	1
Very Good	14	28%	3
Good	15	30%	2
Fair	0	0%	4
Poor	0	0%	4
Total no. Respondents (N)	50	100%	
Mean	4.12		
Verbal Interpretation	**Very Good**		

As shown from the table, out of 50 respondents in verbal interpretation in terms of accessibility. Excellent got the highest frequency of 21 or 42% which is in rank 1, Good with a frequency of 15 or 30%

which is in rank 2, Very Good with a frequency of 14 or 28% which is in rank 3 while Fair and Poor got no verbal interpretation which is rank 4.

Table 3 – The evaluation of the respondents in terms of Accessibility no. 3

Criteria Accessibility	**How would you rate the application's ability in terms of accessing the sent messages, inbox, and drafts by listening to it?**		
Likert Scale	**Frequency f(x)**	**Percentage (p)**	**Rank**
Excellent	16	32%	2
Very Good	22	44%	1
Good	8	16%	3
Fair	4	8%	4
Poor	0	0%	5
Total no. Respondents (N)	**50**	**100%**	
Mean	**4**		
Verbal Interpretation	**Very Good**		

As shown from the table, out of 50 respondents in verbal interpretation in terms of accessibility. Very Good got the highest frequency of 22 or 44% which is in rank 1, Excellent with a frequency of 16 or 32% which is in rank 2, Good with a frequency of 8 or 16% which is in rank 3 while Fair with a frequency of 4 or 8% which is in rank 4 and Poor got no verbal interpretation which is rank 5.

Table 4 – The evaluation of the respondents in terms of Accessibility no. 4

Criteria Accessibility	How would you rate the accessibility of the application in terms of retrieving messages through voice recognition?		
Likert Scale	**Frequency f(x)**	**Percentage (p)**	**Rank**
Excellent	9	18%	3
Very Good	12	24%	2
Good	24	48%	1
Fair	5	10%	4
Poor	0	0%	5
Total no. Respondents (N)	**50**	**100%**	
Mean	**3.5**		
Verbal Interpretation	**Good**		

As shown from the table, out of 50 respondents in verbal interpretation in terms of accessibility. Good got the highest frequency of 24 or 48% which is in rank 1, Very Good with a frequency of 12 or 24% which is in rank 2, Excellent with a frequency of 9 or 18% which is in rank 3 while Fair with a frequency of 5 or 10% which is in rank 4 and Poor got no verbal interpretation which is rank 5.

Table 5 – The evaluation of the respondents in terms of Accessibility no. 5

Criteria Accessibility	How would you rate the accessibility of the application in terms of locating information easily by tapping the screen?		
Likert Scale	**Frequency f(x)**	**Percentage (p)**	**Rank**
Excellent	24	48%	1
Very Good	17	34%	2
Good	9	18%	3
Fair	0	0%	4
Poor	0	0%	4
Total no. Respondents (N)	**50**	**100%**	
Mean	**4.3**		
Verbal Interpretation	**Very Good**		

As shown from the table, out of 50 respondents in verbal interpretation in terms of accessibility. Excellent got the highest frequency of 24 or 48% which is in rank 1, Very Good with a frequency of 17 or 34% which is in rank 2, Good with a frequency of 9 or 18% which is in rank 3 while Fair and Poor got no verbal interpretation which is rank 4.

Table 6 – The evaluation of the respondents in terms of Accuracy no. 1

Criteria Accuracy	How would you rate the accuracy of every data and information produced by the application?		
Likert Scale	**Frequency f(x)**	**Percentage (p)**	**Rank**
Excellent	3	6%	4
Very Good	15	30%	2
Good	24	48%	1
Fair	8	16%	3
Poor	0	0%	5
Total no. Respondents (N)	**50**	**100%**	
Mean	**3.26**		
Verbal Interpretation	**Good**		

As shown from the table, out of 50 respondents in verbal interpretation in terms of accuracy. Good got the highest frequency of 24 or 48% which is in rank 1, Very Good with a frequency of 15 or 30% which is in rank 2, Fair with a frequency of 8 or 16% which is in rank 3 while Excellent with a frequency of 3 or 6% which is in rank 4 and Poor got no verbal interpretation which is rank 5.

Table 7 – The evaluation of the respondents in terms of Accuracy no. 2

Criteria Accuracy	How would you rate the accuracy of the voice navigation guide (VNG) in guiding the user to use the application?		
Likert Scale	**Frequency f(x)**	**Percentage (p)**	**Rank**
Excellent	11	22%	2

Very Good	24	48%	1
Good	8	16%	3
Fair	5	10%	4
Poor	2	4%	5
Total no. Respondents (N)	**50**	**100%**	
Mean	**3.74**		
Verbal Interpreter	**Good**		

As shown from the table, out of 50 respondents in verbal interpretation in terms of accuracy. Very Good got the highest frequency of 24 or 48% which is in rank 1, Excellent with a frequency of 11 or 22% which is in rank 2, Good with a frequency of 8 or 16% which is in rank 3 while Fair with a frequency of 5 or 10% which is in rank 4 and Poor with a frequency of 2 or 4% which is in rank 5.

Table 8 – The evaluation of the respondents in terms of Accuracy no. 3

Criteria Accuracy	How would you rate the accuracy of the words and phrases decoded by the application in terms of receiving voice input from the user?		
Likert Scale	**Frequency f(x)**	**Percentage (p)**	**Rank**
Excellent	16	32%	2
Very Good	25	50%	1
Good	7	14%	3
Fair	2	4%	4
Poor	0	0%	5
Total no. Respondents (N)	**50**	**100%**	
Mean	**4.1**		
Verbal Interpretation	**Very Good**		

As shown from the table, out of 50 respondents in verbal interpretation in terms of accuracy. Very Good got the highest frequency of 25 or 50% which is in rank 1, Excellent with a frequency of 16 or 32% which is in rank 2, Good with a frequency of 7 or 14% which is in rank 3 while Fair with a frequency of 2 or 4% which is in rank 4 and Poor got no verbal interpretation which is rank 5.

Table 9 – The evaluation of the respondents in terms of Accuracy no. 4

Criteria Accuracy	**How would you rate the accuracy of the pronunciation or diction of the speech recognition in terms of providing voice output back to the user?**		
Likert Scale	**Frequency f(x)**	**Percentage (p)**	**Rank**
Excellent	23	46%	1
Very Good	15	30%	2
Good	12	24%	3
Fair	0	0%	4

Poor	0	0%	4
Total no. Respondents (N)	**50**	**100%**	
Mean	**4.22**		
Verbal Interpretation	**Very Good**		

As shown from the table, out of 50 respondents in verbal interpretation in terms of accuracy. Excellent got the highest frequency of 23 or 46% which is in rank 1, Very Good with a frequency of 15 or 30% which is in rank 2, Good with a frequency of 12 or 24% which is in rank 3 while Fair and Poor got no verbal interpretation which is rank 4.

V. CONCLUSION AND FURTHER RESEARCH

Conclusion

Based on the results of the surveys, conclusions were drawn:

1. Majority of the respondents were Male.
2. Majority of the respondents were in their adultery period ranging ages 32-41.
3. Majority of the respondents were working as a massage therapist.
4. Majority of the respondents in terms of highest educational attainment took vocational courses or training on special course.
5. Majority of the respondents rated the overall function of the application Excellent.
6. Majority of the respondents rated the security of the application Secured.

Recommendation

Based on the results from the surveys of the android based voice recognition e-mail application, the proponents can say that the proposed application is given highly recommended ratings by the respondents which are the blind or visually impaired persons. The respondents were very overwhelmed upon using the application for they felt cared and loved knowing that there are some developers who really want to develop new trend in computer technology for people like them. But there are some suggestions and feedbacks received from the respondents

Recommendations from the feedbacks are the following:

1. The application should understand other languages.
2. The application should have other features wherein the user may use zero load balance in their mobile phone.
3. The application should be also applicable on IOS.

Therefore, it is highly recommended that the application will continuously be modified to best fit the needs of the said disabled people.

REFERENCES

1. Griffin-Shirley N, Banda DR, Ajuwon PM, et al. A Survey on the Use of Mobile Applications for People who Are Visually Impaired. Journal of Visual Impairment & Blindness. 2017;111(4):307-323. doi:10.1177/0145482X1711100402
2. Abhishek Thakur, Rajesh Kumar, Naveen Kumar, (Dept. of ECE, IGCE, Abhipur, Mohali, Punjab, India, "Automatic Speech Recognition System for Hindi Utterances with Regional Indian Accents" https://pdfs.semanticscholar.org/7117/6b84ca0ec49b6c9f7120d35469ba6e7b8859.pdf
3. Jerrick Chua, Unisse chua, Cessar de Padua, Janelle Isis Tan, and Mr. Danny Cheng who studies college of Computer Studies at De La sale University of Manila, "A Filipino Hands-free Text Messaging Application"
4. https://pdfs.semanticscholar.org/2127/7e621115cbadc074bdbbf9538bcbf090dae5.pdf?_ga=2.94376604.1890603161.1552798848-910769576.1552798848
5. Ibrahim, Jafar Ali S., and M. Thangamani. "On-Line Patient Interaction Method for Disease Discovery by Data Mining Classification." *Journal of Computational and Theoretical Nanoscience* 16, no. 2 (2019): 573-575.
6. Elayaraja, D, Jafar Ali Ibrahim S et al. "Design Parametric Optimization Of Wall Following Robot." *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12, no. 8 (2021): 2072-2080.

7. Ibrahim, S. Jafar Ali, and M. Thangamani. "Enhanced singular value decomposition for prediction of drugs and diseases with Hepatocellular carcinoma based on Multi-Source Bat Algorithm based random walk." *Measurement* 141 (2019): 176-183.
8. Karthikeyan, B., K. Alhaf Malik, D. Bujji Babbu, K. Nithya, S. Jafar Ali Ibrahim, and NS Kalyan Chakravarthy. "Survey of Cooperative Routing Algorithms in Wireless Sensor Networks." *Annals of RSCB* 25, no. 1 (2021): 5316-5320.
9. Miguel Carlos Pascual, John LlyodLagura, Rehum Jay Rabadan, and AlexanderTabac from electronics and engineering graduates of De La Salle University (DLSU), "SMS for Blind Using Braille" https://aboutphilippines.org/Video-descriptions/SMS-for-the-Blind.pdf