Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

ERFOA Technique – An Important Tool For The Visually Impaired

S. Murugesan[1]*, Dr. N. Balajiraja[2]

*[1]*Research Scholar, JJ College of Arts and Science, Pudukkottai, Affilated to Bharathidaasan University, Tamil Nadu, India*
kpsmurugesan@gmail.com
[2]Assistant Professor, PG and Research, Department of Computer Science, J.J. College of Arts and Science (Autonomous), Pudukkottai, Affilated to Bharathidaasan University, Tamil Nadu, India nbalajiraja@gmail.com

VIIs (Visually Impaired People) may not be able to read books, newspapers, and magazines because the words and letters are hard for them to see. Blindness affects millions of people. Despite the fact that inclusiveness is a progressive development guideline, VIIs have historically been on the margins of society and, more especially, of higher education. Because of this, it is crucial to make sure they are fully included in all aspects of life, with particular care for them. The goal of this work is to help VIIs by showing them how to use the internet to fill out online application forms. The goals of the proposed approach are to search, filter, and fill out online application forms utilizing voice interactions and Machine Learning Techniques (MLTs).

Keywords: Visually Impaired, Grouping, Binary Encoding, Evaluation, Feature selection.

Introduction: It is estimated that more than 600 million Individuals are visually impaired in the world [1]. They have been treated in society as the marginalized. Researches and studies have also focused on enhancing their lives by introducing gadgets with sensors for reducing their problems due to sight. Assisting the blind started in 1800s when Charles Barbier from Napoleon Bonaparte's army developed a coded language called "night writing" which was used in communications in darkness and when soldiers carrying lanterns were killed. His work was carried on by the 11 year old blind French, Louis Braille. Figure 1 depicts Night Writing

Fig 1 – Night Writing

Motivated, Braille enrolled in Paris's National Institute of the Blind and created a method using only six dots per cell. His technique gained widespread recognition as the de facto standard for blind people's written communication. Figure 2 shows the dots used by Braille.

Fig. 2 – Braille System of Dots

Though Braille's system with modifications is predominantly used by VIIs, it is difficult for them to perceive external environment. VIIs who have poor vision, which prevents them from seeing even regular written paper since only large letters are visible, seem to benefit little from glasses. The modern technological environment, where everything is digital, has the potential to advance these VIIs. Incorporating VIIs into a nation's economy may be greatly aided by IT. Instead of utilizing conventional analog voice carrying technologies, VOIP (Voice Over Internet Protocol) enables digital voice communication over broadband Internet connections. With their Voice-controlled Intelligent Personal Assistants (VCIPAs), large corporations like as Amazon and Google have created a new paradigm of interaction for the blind and visually impaired. VIIs may utilize these conversational interfaces for a variety of purposes. Nevertheless, research indicates that VII-assistance devices are being employed broadly rather than specifically for VIIs [2][3].

VIIs have a significant role in society as well. Furthermore, Dr. Bjorn has hinted that social or educational relationships may suffer as a result of visual impairment. It may have an impact on a person's social, professional, and natural development of intellect and academic aptitude [4]. The majority of VIIs are capable of voice-guided digital device operation. Computers can be used more easily by VIIs thanks to screen readers with speech capability, like as Jaws and NVDA. Therefore, the implementation of an assistive system to enable VIIs to access websites for online enrollment or employment applications is the main emphasis of this study. VIIs

may benefit from the suggested approach if they want to use their voice to access online application forms. This study is structured as a review of the literature on VIIs that comes after this section. The suggested ERFOA (Extracting Required Fields of Online Applications) Technique is explained in the third part, and section four displays the outcomes of implementations. This paper's conclusion is found in Section 5.

Reviewing Accompanying Works: At least fifteen million people of Indian descent call India home. A variety of VII devices are accessible, like as canes or guiding dogs. Recent lightning-fast developments in hardware and software have led to the rise in popularity of a plethora of smart devices that claim to aid VIIs. To aid VIIs in dealing with elevator-related problems, a number of methods were developed. This kind of technology was initially introduced in 2014. Built from an intelligent lift that responds to voice commands, it facilitated human-machine communication by translating spoken commands into actions [5]. It helped the crippled and paralyzed despite its complex electronics and average speech recognition performance. In2015, a second system of this kind was developed to assist VIIs [6]. There were a lot of expenses, a poor speech recognition accuracy, and only 10 voice commands in the system. Without voice confirmations, the third VII system from 2016 showed low power consumption, cheapness, and great efficiency in speech recognition [7]. Technologies that are accessible by voice have been used often. One typical method of facilitating accessibility is the use of speech input and output. Screen readers (such as Apple's [8] or JAWS [9]) provide audio output for users with visual impairments, while voice dictation software Dragon [10] is one of the most popular kinds of accessible speech interaction. For those who have trouble moving their hands or feet, speech input has become useful for many things, including as text entry on computers and mobile phones, operating wheelchairs, and even "free-hand" artwork. Due to the efficiency it provides for text entry and browsing, voice input on mobile devices is more widespread for users with visual impairments than for sighted users [15][16]. Users with various forms of disability have also been the subject of research on speech interaction. People who have trouble speaking may benefit from speech treatment [30] and practice [18] using electronic speech interaction. Short, directive sentences that were often imprecise were formulated by individuals with speech and physical disabilities in an exploratory research on how they generate instructions for a voice interface [19]. Research on speech interaction for people with cognitive impairments has also yielded some useful results, such as design recommendations for spoken conversation aides for dementia users [20] and audio prompts to help with everyday chores [21]. Unlike the aforementioned work, our main emphasis is on conversational interaction using simply speech and a generic, off-the-shelf IPA. A plethora of approaches have been developed to address the needs of those who are visually impaired. After reviewing several studies in pattern recognition, the authors of article [22] provide a practical model for object detection and gesture analysis utilizing computer vision and machine learning for decision making. In order to identify objects in photos prior to categorization using NNs, a network model was used in the aforementioned article [23]. In [24], a text reader was introduced that could scan movies and photos and pull text from digital comics. Feature and word extractions, as well as pre-processing and segmentations, were used. Then, it used POS (Parts of Speech) to label categories. In [25], the authors examined the difficulties of using

conversational assistants on smartphones for mundane activities such as creating reminders. The trust and intelligence level of intelligent personal assistants, such as Google Home and Amazon Echo, were investigated in the research conducted by [26].

PROPOSED SYSTEM ERFOA: VIIs struggle from a variety of issues in their daily life. Studies on vision in humans and other animals have shown that visual perception, while a crucial tool for navigating the environment, may be substituted. Voice assistants are smartphone apps that aid VIIs. With the use of a speaker, the user may communicate, send messages, view notes, and make phone calls. Many years have been spent researching assistive devices for VIIs [27]. The proposed ERFOA waits for user input in the form of speech and uses the ERFOA-SR Algorithm to identify text from voice. After that, the text is utilized as a query to look for the desired website on the internet. Once again, the link is used to extract all of the form's fields (feature extraction). Only input fields are chosen, and features are chosen according to a weight that is supplied. After the user demands the information for each field, it is transformed to text, filled out, and finally submitted. The ERFOA architecture is shown in Figure 3, and the ERFOA-SR Algorithm is described below.

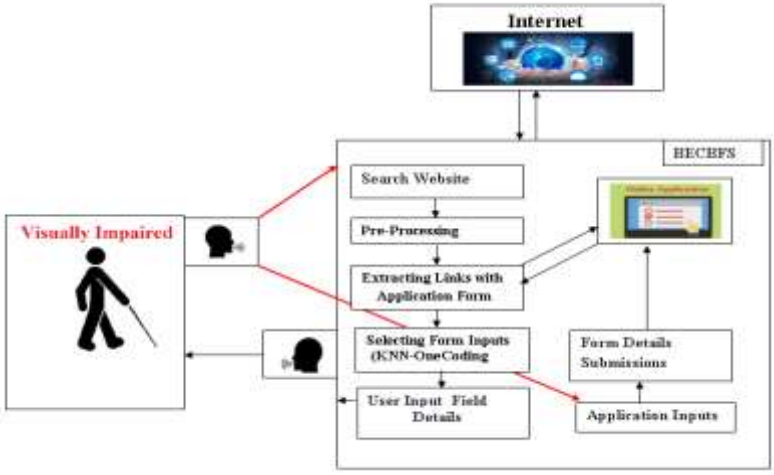

Fig. 3 – ERFOA Architecture

ERFOA-SR Algorithm //Speech Recognition Algorithm
Input: Analog Signals
Outtext=””
#Initialize Microphone Parameters
Set_ speech_speed= 125
Set_Volume_Max=1 # Set volume 0-1
Set_Noise_Level =Adjust ambient_noise
While Input Voice =none
For Each Sound in decibels (dB)
Slice into 20 Milli second Acoustic Frames

for each acoustic frame of 20 to 40 milliseconds
converter sound to digital format
Recognized Text
End For
Outtext=outtext + Recognized Text
End While
Output: Spoken Text

ERFOA Pre-Processing: Once the website is identified in the internet the home page is downloaded with all its links. Figure 4 depicts the comparative webpage fetch time.

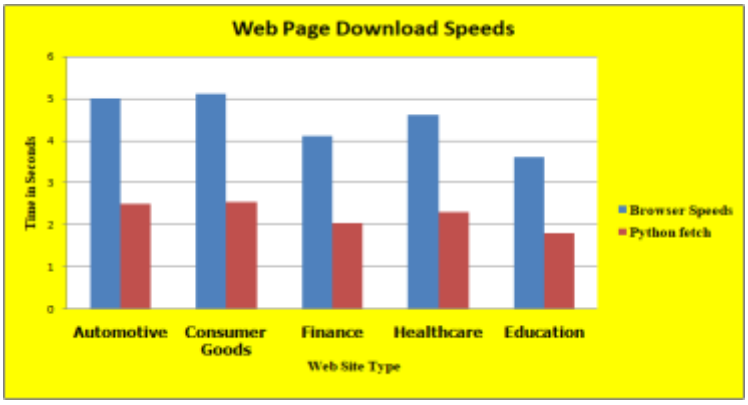

Fig. 4 – Web Page Download Speeds

Pre-processing is a necessary step in data processing. It is also a data cleaning/preparation process. In ERFOA, since the user is a VII, it not possible for him/her to choose the online applications forms in a website. Hence, pre-processing in this work is filtering unwanted links (Filtering) and retaining only the links that lead to the online application page/Form. Figure 5 depicts the basic components of the webpage found in the study

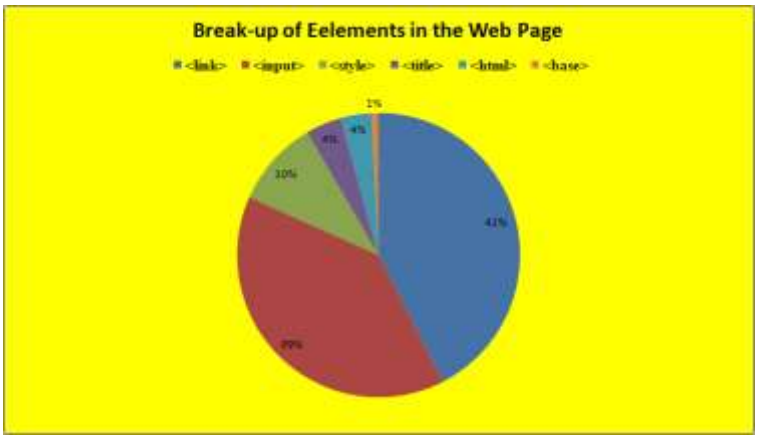

Fig. 5 – Break-up of Web Page Elements

This work uses ERFOA-Filters. Once, the site is found all the links in home page are fetched by the applications where only links which lead to the application form of the queried website is chosen. ERFOA-Filtering Algorithm is listed below

ERFOA-Filtering Algorithm

```
Input: Href Links (Google) //Tracing Application Form Websites
OutLink=""
While OutLink=none
    For Each Http 1.. N //N is the Number of Links returned by Google
          InLink=Http
           For Each Word in InLink
              If Word="Application" then
                 OutLink=InLink
                 Exit For
            End For
   End For
End While
Output: OutLink
```

ERFOA Feature Extraction: People are often required to complete online forms. In America, for instance, it is estimated that 70 million professionals apply online, often completing the same form. Presenting a user interface that can be utilized to quickly and precisely collect data from several users is the aim of a form. Consequently, a form often includes a number of standard UI components, all of which are usually referred to as fields, including text boxes, drop-down menus, and date pickers. It is anticipated that the user will fill up these fields. Usually, a form's fields are completed in a certain sequence. This offers a chance to increase production via the use of predictive algorithms. When predicting the value

of the field the user is presently filling out, existing methods do not take into account the values of other fields on the form. Due to their poor vision, these methods are unable to simulate the naturally occurring dependencies between fields, such as those between a field meant to record the name of the form filler and another field meant to record the filler's address. As a result, ERFOA's feature extraction is limited to extracting all of the fields from the form data by selecting every field in the form. The suggested approach assumes that the application form that is fetched is Fm with a collection of tags Ftgi such that i={0,1,2….n} where n is the number of fields. Equation (1) may therefore be used to express the conditional probability of selecting every tag that makes up inputs.

$$P(Ftg_i) = Ftg_i \text{ context} / P(\text{context}(Ftg_i)) \quad \text{.....(1)}$$

The labels which need inputs are selected by ERFOA-IF Algorithm.

ERFOA-IF Algorithm

```
Input: Application Form Tags
Output: Text Fields which Require Inputs
# extract all form details
#Initialize Details
data={}
ctr1=0
ctr2=0
    for Each tag in form
        ctr1=ctr1+1
        if input_tag["type"]="input"
            if input_tag["type"] == "hidden":
                ctr2=ctr2+1
                data[input_tag["name"]] = input_tag["Default value"]
            else
                ctr2=ctr2+1
                data[input_tag["name"]] = ""
          end if
 For loop End
Output :  data[]
```

These selected features are then passed on **ERFOA-SR Algorithm** for Voice to Text operations for interactively filling the application form. Figure 6 depicts the input processing of selected feature inputs.

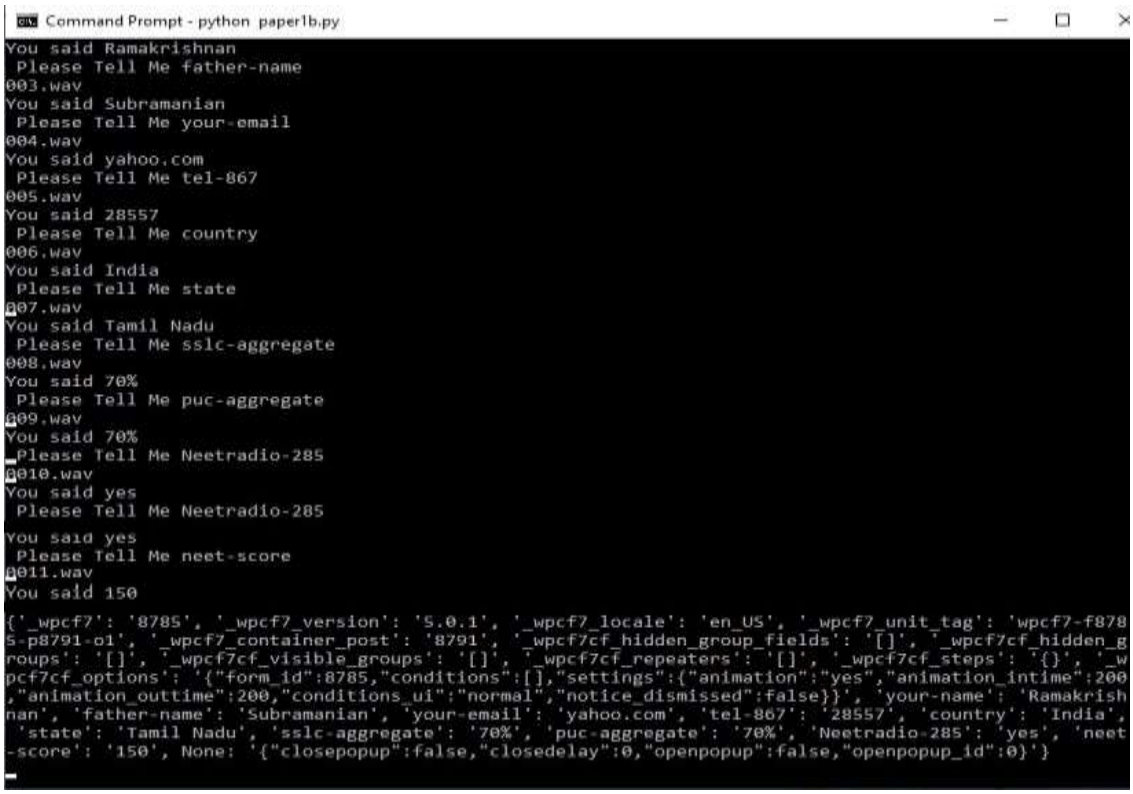

Fig. 6 – Inputs of Selected Features and Form Submission

Experimental Results: The proposed ERFOA was implemented using python 3 running on windows 10 with an AMD processor. The stage wise results are shown as figures. The feature selections were evaluated with KNN Random forest and Decision Trees Algorithms for assessing its performance where all the three classifiers used the same inputs. The VII utters the name of a website which is then received by the system and the website is searched for in the internet. Figure 7 shows the VOIP inputs and Outputs with the site searched.

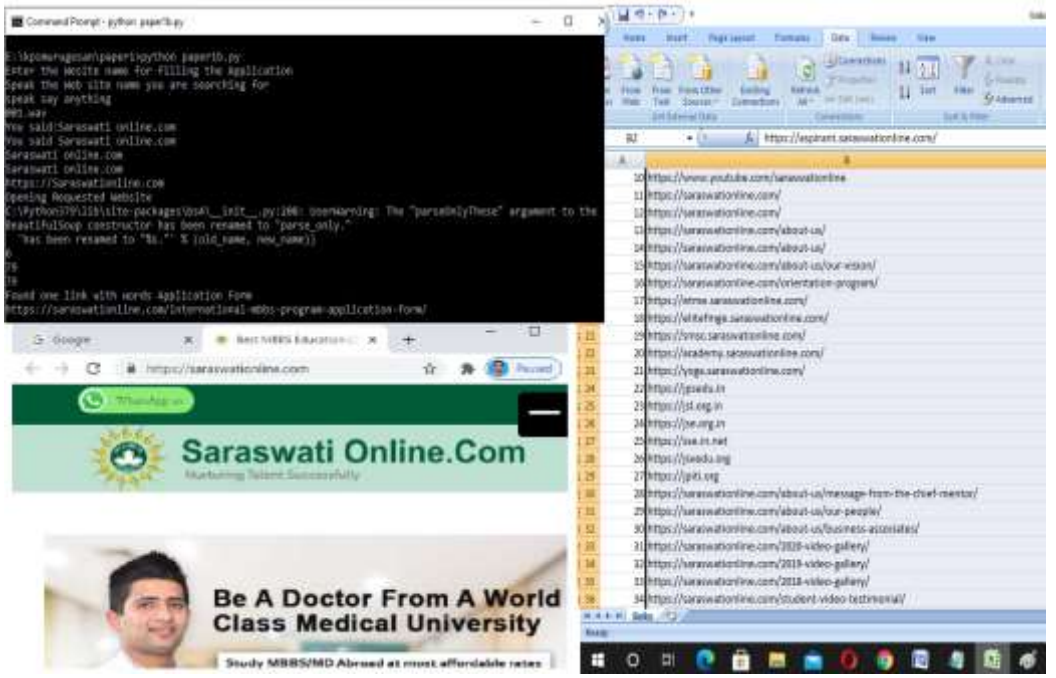

Fig. 7 – ERFOA Pre-processing Output

It can be seen from the above figure that the site was searched which had 79 links and only one link existed for its online application. The links were removed from further processing. In Feature extraction all the HTML tags were fetched by the proposed technique for its feature selection. All the inputs tags from the application form are selected in the beginning as depicted in Figure 8.

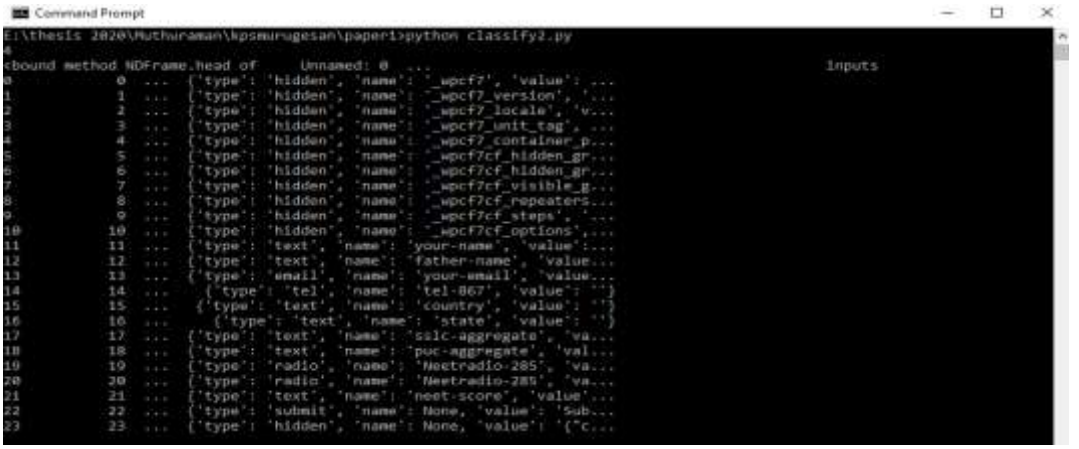

Fig. 8 – ERFOA Selected Features

Model Evaluation: Sample data (also known as "training data") is used by the majority of machine learning algorithms to extract features. Without being specifically programmed to do so, a mathematical model is developed based on these characteristics and utilized to make judgments or predictions. The model was evaluated using one-hot encoding to assess the chosen features. Every single bit of data is represented in a form that a computer can comprehend using one-hot encoding. With one-hot encoding, every element in a vector is zero, with the exception of those vectors that have a value of one. As a result, labels are chosen according to their numerical weights based on their significance. In order to ascertain a machine's cost status, one-hot encoding is used. Because it just requires two flip-flops, changing a machine's found state is almost as quick. It is very simple to put into practice. Table 1 describes how digits 0 through 7 are represented in binary, gray code, and one-hot:

Table 1 – One Hot Coding Example

Decimal	**Binary**	**Gray code**	**One-Hot**
0	0	0	0
1	1	1	1
2	10	11	10
3	11	10	100
4	100	110	1000
5	101	111	10000
6	110	101	100000
7	111	100	1000000

The selected features are then encoded with binary values for adding weights and create a dynamic dataset for classification algorithms to learn. Figure 9 depicts the output of Encoding and creation of the Dynamic dataset.

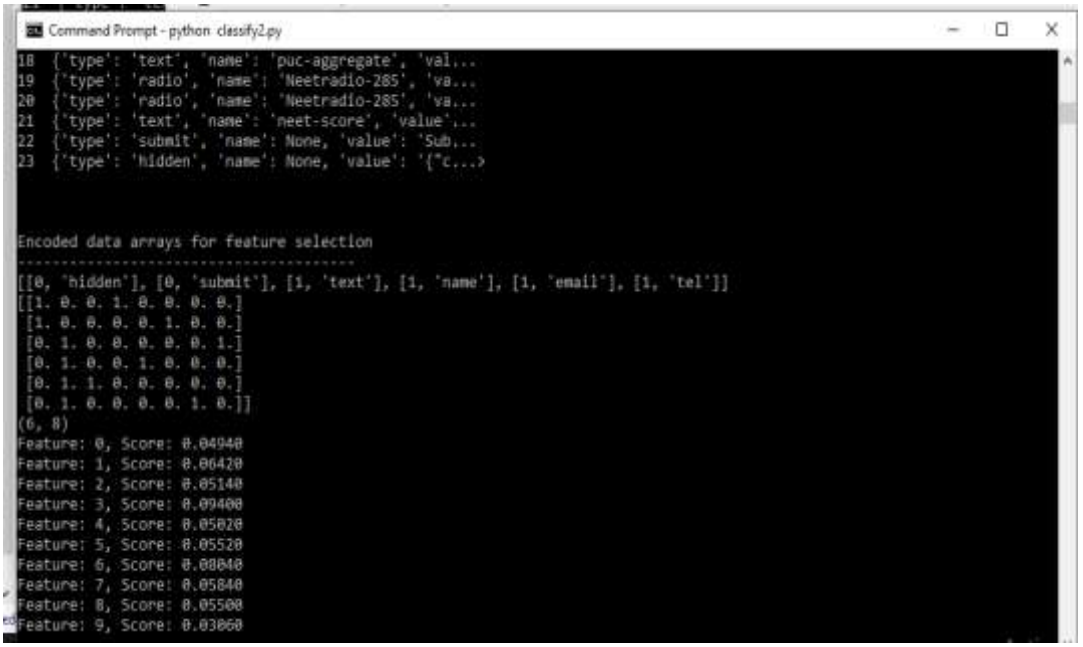

Fig. 9 – ERFOA One-Hot Encoding

Since the goal is to construct a better model, tracking how well a model performs on a validation set is a great approach to get feedback on how a suggested model is functioning. All of the classifiers used in the ERFOA-IF Algorithm feature selection process, including DT (Decision Trees), RF (Random Forest), and KNN, were part of the evaluation of the suggested system. DT is a supervised ML method that works for regression and classification tasks alike. A decision tree is only a set of steps taken in a certain order to accomplish a goal. RF is a method for machine learning that uses trees to make judgments; it takes use of the strength of many trees. In order to get the ultimate result, it merges the output of several Decision Trees that are generated at random. On display in Figure 10 are the outcomes of the ERFOA selection scores assessment. The suggested approach uses the most characteristics for VIIs' input text fields, which is the maximum.

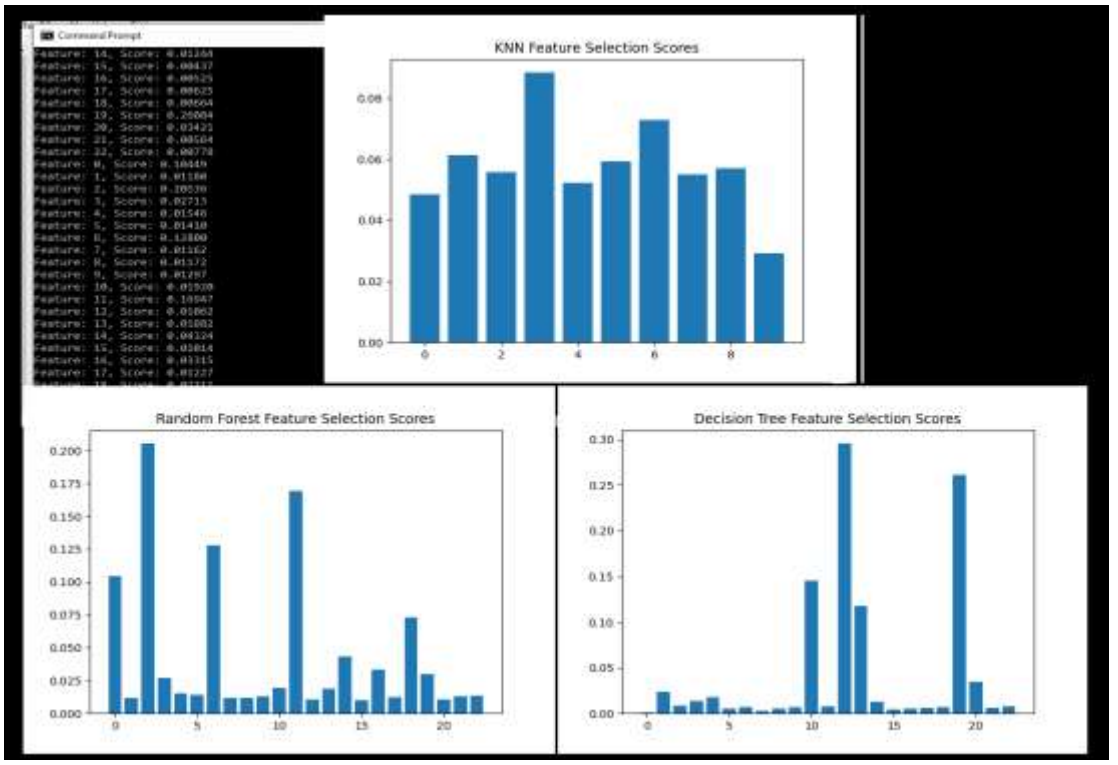

Fig. 10 – Performance Evaluation of ERFOA-IF Algorithm selected features using Classifiers

Conclusions and future work

Words in information sources cannot be recognized by VIIs. Blindness affects millions of people. There are at least 15 million Indians at home. Aids such as walking canes and guiding dogs are available for those with visual impairments. A plethora of intelligent gadgets that can assist VIIs have recently emerged, thanks to the fast development of both software and technology. Despite the fact that inclusiveness is a progressive development guideline, VIIs have historically been on the margins of society and, more especially, of higher education. When applying online, you'll be required to fill out each and every area on the form with your own information. There is usually a predetermined and shown sequence for filling out the fields on a form. Since the point of a form is to provide a means for users to submit information, it stands to reason that a form filling prediction model would aim to shorten the time it takes for users to complete the form. Feature selections are therefore crucial to the aforementioned goal. Selecting characteristics based on their relative importance to a target prediction variable or output is known as feature selection. This article has developed and launched a free tool to help people with disabilities apply for employment online using mobile phones or the internet. It follows that current systems are amenable to the suggested method, which will ultimately help VIIs and people everywhere.

References

[1] Global Data On Visual Impairment, World Health Organization, Https://Www.Who.Int/Blindness/Publications/Globaldata/En/.
[2] Stefania Druga, Randi Williams, Cynthia Breazeal, And Mitchel Resnick. 2017. "Hey Google Is It Ok If I Eat You ?" Initial Explorations In Child-Agent Interaction. In Proceedings Of The 2017 Conference On Interaction Design And Children (Idc '17), 595–600. Https://Doi.Org/10.1145/3078072.3084330
[3] Amanda Purington, Jessie G. Taft, Shruti Sannon, Natalya N. Bazarova, And Samuel Hardman Taylor. 2017. "Alexa Is My New Bff": Social Roles, User Satisfaction, And Personification. In Proceedings Of The 2017 Chi Conference Extended Abstracts On Human Factors In Computing Systems (Chi Ea '17), 2853– 2859. Https://Doi.Org/10.1145/3027063.3053246
[4] Hersh, M., And Johnson, M. A. Eds., "Assistive Technology For Visually Impaired And Blind Individuals", Springer Science & Business Media, 2010.
[5] Sooxma Technologies, "Voice Operated Intelligent Lift/Elevator," My College Projects, 2014.
[6] J. F. A. O. L. Å. R. Oskar Edwertz, Smartlift: Voice-Activated, Gothenburg: Chalmers University Of Technology, 2015.
[7] S. M. J. B. Vrajeshprajapati, "Voice Initialized Elevator," International Journal Of Electrical And Electronics Research, Vol. 4, No. 1, Pp. 154-157, 2016.
[8] Voiceover For Mac: Voiceover Overview. Retrieved September 14, 2017 From Https://Support.Apple.Com/Kb/Ph22549?Locale=En_Us
[9] Jaws Screen Reader - Best In Class. Retrieved September 14, 2017 From Http://Www.Freedomscientific.Com/Products/Blindness/
Jaws
[10] Dragon Naturallyspeaking - World's Best-Selling Speech Recognition Software | Nuance. Retrieved September 14, 2017 From Https://Www.Nuance.Com/Dragon.Html
[11] Bill Manaris, Valanne Macgyvers, And Michail Lagoudakis. 2002. A Listening Keyboard For Users With Motor Impairments — A Usability Study ∗. International Journal Of Speech Technology 5, 4: 371– 388.
[12] Amber Wagner, Ramaraju Rudraraju, Srinivasa Datla, Avishek Banerjee, Mandar Sudame, And Jeff Gray. 2012. Programming By Voice: A Hands-Free Approach For Motorically Challenged Children. In Proceedings Of The 2012 Acm Annual Conference Extended Abstracts On Human Factors In Computing Systems Extended Abstracts (Chi Ea '12), 2087–2092. Https://Doi.Org/10.1145/2212776.2223757
[13] Maia Naftali And Leah Findlater. 2014. Accessibility In Context: Understanding The Truly Mobile Experience Of Smartphone Users With Motor Impairments. In Proceedings Of The 12th International Acm Sigaccess Conference On Computers And Accessibility (Assets '14), 209–216. Https://Doi.Org/10.1145/2661334.2661372
[14] Susumu Harada, Jacob O. Wobbrock, Jonathan Malkin, Jeff A. Bilmes, And James A. Landay. 2009. Longitudinal Study Of Individuals Learning To Use
[15] Shiri Azenkot And Nicole B. Lee. 2013. Exploring The Use Of Speech Input By Blind Individuals On Mobile Devices. In Proceedings Of The 15th International Acm Sigaccess Conference On Computers And Accessibility (Assets '13), 1–8. Https://Doi.Org/10.1145/2513383.2513440
[16] Hanlu Ye, Meethu Malu, Uran Oh, And Leah Findlater. 2014. Current And Future Mobile And Wearable Device Use By Individuals With Visual Impairments. In Proceedings Of The Sigchi Conference On Human Factors In Computing Systems (Chi '14), 3123–3132. Https://Doi.Org/10.1145/2556288.2557085

[17] Vikas Ashok, Yevgen Borodin, Svetlana Stoyanchev, Yuri Puzis, And I. V. Ramakrishnan. 2014. Wizard-Of-Oz Evaluation Of Speech-Driven Web Browsing Interface For Individuals With Vision Impairments. In Proceedings Of The 11th Web For All Conference (W4a '14), 1–9. Https://Doi.Org/10.1145/2596695.2596699

[18] Susan Koch Fager. 2017. Speech Recognition As A Practice Tool For Dysarthria. Seminars In Speech And Language 38, 3: 220–228. Https://Doi.Org/10.1055/S-0037-1602841

[19] Jan Derboven, Jonathan Huyghe, And Dirk De Grooff. 2014. Designing Voice Interaction For Individuals With Physical And Speech Impairments. In Proceedings Of The 8th Nordic Conference On Human-Computer Interaction: Fun, Fast, Foundational (Nordichi '14), 217–226. Https://Doi.Org/10.1145/2639189.2639252

[20] Maria Klara Wolters, Fiona Kelly, And Jonathan Kilgour. 2016. Designing A Spoken Dialogue Interface To An Intelligent Cognitive Assistant For Individuals With Dementia. Health Informatics Journal 22, 4: 854–866. Https://Doi.Org/10.1177/1460458215593329.

[21] Clare Carroll, Catherine Chiodo, Xin Adena Lin, Meg Nidever, And Jayanth Prathipati. 2017. Robin: Enabling Independence For Individuals With Cognitive Disabilities Using Voice Assistive Technology. In Proceedings Of The 2017 Chi Conference Extended

[22] Aditya Raj, Manish Kannaujiya, Ajeet Bharti, Rahul Prasad, Namrata Singh, Ishan Bhardwaj " Model For Object Detection Using Computer Vision And Machine Learning For Decision Making " International Journal Of Computer Applications (0975 – 8887) Volume 181 – No. 43, March 2019.

[23] Selman Tosun, Enis Karaarslan "Real-Time Object Detection Application For Visually Impaired Individuals: Third Eye". Ieee Conferences 2018.

[24] "Jayshree R Pansare, Aditi Gaikwad, Vaishnavi Ankam, Priyanka Karne And Shikha Sharma "Real - Time Text Reader" International Journal Of Computer Applications 182(34):42-45, December 2018.

[25] Aarthi Easwara Moorthy And Kim-Phuong L Vu. 2015. Privacy Concerns For Use Of Voice Activated Personal Assistant In The Public Space. International Journal Of Human-Computer Interaction 31, 4: 307–335. Https://Doi.Org/10.1080/10447318.2014.986642

[26] Stefania Druga, Randi Williams, Cynthia Breazeal, And Mitchel Resnick. 2017. "Hey Google Is It Ok If I Eat You ?" Initial Explorations In Child-Agent Interaction. In Proceedings Of The 2017 Conference On Interaction Design And Children (Idc '17), 595–600. Https://Doi.Org/10.1145/3078072.3084330

[27] R. Mengko And A. Ayuningtyas."Indonesian Tts System Using Syllable Concatenation: Speech Optimization" Proc. International Conference On Instrumentation, Communication, Information Technology, And Biomedical Engineering (Icici-Bme), November 2013, Pp. 412-415.

[28] Viola, P. And Narasimhan, M. Learning To Extract Information From Semi-Structured Text Using A Discriminative Context Free Grammar. Proceedings Of The 28th Annual International Acm Sigir Conference On Research And Development In Information Retrieval, 2005, Association For Computing Machinery, 330-337.