ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/349845098

Using NLP Techniques for Enhancing Augmentative and Alternative Communication Applications

Research · January 2021

DOI: 10.13140/RG.2.2.14855.32161/1

CITATION
1

READS
544

1 author:

Hardik Srivastava

5 PUBLICATIONS **3** CITATIONS

SEE PROFILE

All content following this page was uploaded by Hardik Srivastava on 09 November 2022.

The user has requested enhancement of the downloaded file.

International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-06, Issue-10, JAN 2021

Using NLP Techniques for Enhancing Augmentative and Alternative Communication Applications

Hardik Srivastava

Department of Computer Science and Engineering, SRM Institute of Science and Technology, India.

Abstract - According to the WHO, around one billion people suffer from disorders in speaking and experience motor-skill challenges which can often lead to exclusion from the society. An enhanced and accelerated device for communicating is something the disabled people cannot afford, who still form the world's largest minority to experience discrimination. So, in this paper I would like to propose an application, HearMeOut, that uses artificial intelligence techniques to implement state-of-the-art algorithms and Natural Language Processing for predicting meaningful and appropriate text, generating sentences using pictograms and facilitating short communication using speech recognition. These specifications are delivered to the user via a language assistant called LGenie (Language Generation through Input Event). This will help reduce the communication barrier between the speechless and their interlocutors thus allowing them to integrate into the existing social structures and make the world a more inclusive place.

Index Terms - Augmentative and Alternative Communication, NLP, Transfer Learning, Sign to Text Prediction, Speech Recognition.

I. INTRODUCTION

Communication is what differentiates the normal from deaf/mute. The only way for them to communicate is through sign language or through an interpreter. It is possible to build prosthetic devices for such users by linking a suitable physical interface with a speech synthesizer. However, while in normal conversations the speech rates are around 150-200 words per minute and skilled typists can achieve a maximum rate of 30-40, there are conditions which impair physical ability to speak. These cause more general loss of motor function and typically speech prosthesis users can only output at best 10-15 wpm using a keyboard. This prevents natural conversation due to delays in turn-taking. Though abled people do not meet any reason to learn sign language yet it is very important to the disabled for communication with both the normal people and among themselves. However, we do not understand its importance and keep giving little attention to it and tend to ignore its importance. A cost-effective and handy solution which helps the deaf/mute and normal people communicate normally is the need of the hour. Applications like Augmentative and Alternative Communication (AAC) have emerged as adaptive solutions in enabling effective communication for people with disabilities. AAC modules facilitate communication in 2 methods. First is where the user builds a sentence by selecting subsets of suggested words and then synthesizing them. The other method is through pictograms where a user expresses his thoughts and feelings via pictures. HearMeOut uses a word level sign language dataset to classify the most probable words from the gestures of the impaired person captured using an external/internal camera and then transforming them into the most meaningful and probable sentence using state-of-the-art algorithms. It also comes with a pictogram-based system and a speech recognition module which works on Trigger Word Detection to facilitate communication. The idea is to somehow mimic the traditional communication boards and expand them into mobile computing context. Considering the limitations, the pictograms are usually organized in categories and are accessible through hierarchies of items. Using NLP, presenting the most appropriate choices using the least number of steps or screens to navigate is what the proposed solution is capable of. When talking about the security

of the application, the structure of HearMeOut is meant to be designed in such a way that it has multiple sources of intelligence, such that failure in any single part of the system will not destroy the efficacy of the entire system. There is absolutely no storage of user data which minimizes the chances of any data leakage. It is a continuous process of input and output sans storage.

 DOI : 10.35291/2454-9150.2021.0031 © 2021, IJREAM All Rights Reserved.

II. Materials and Methods

The proposed HearMeOut application is a solution that translates the concept of communication boards and interpreters to a mobile environment for people with impaired hearing or speech or suffering from any motor-skill challenges by helping them break the communication barrier and mimic natural conversation with an abled person. It is built using a research technique known as Transfer Learning by implementing the ResNetV2, used for classifying the gestures or signs made by a deaf/mute person and predicting the most probable words assigned to the respective images using a word level sign language dataset. These words will be fed into multiple algorithms where after deep pre-processing, the most meaningful and probable sentence will appear on the screen just below the video feed of the disabled person. The person will have to use sign language to communicate with the system and it will automatically keep generating English text of what those signs actually mean. A text-to-speech (TTS) module will narrate the entire sentence if a pause of more than 7 seconds is observed while capturing the frames of the person communicating. For a person suffering from motor-skill challenges, a pictogram-based communication system can be chosen as the medium. I have shown in Fig. 1 a scenario where the disabled person wishes to eat a mango. To express this desire, he/she has to follow certain steps, each one corresponding to choosing an item on the screen. The selections are highlighted with a contrasting colour.

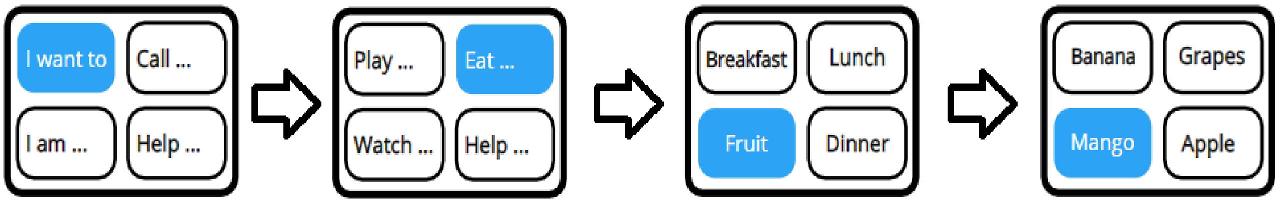

Fig 1. Scenario showing the sequence of steps to eat a mango

In Fig. 1 “I want to ...” is followed by the next prompt “Eat ...”, followed by “Fruit ...”, and finally “Mango”. In this way, simple sentences based on your current location and time of day, will be prompted thus allowing the person to communicate with his/her interlocutors with minimum effort thereby facilitating contextual communication. Although trivial for an individual with good motor skills, this sequence of decisions can typically be slow or troublesome for people with severe motor-skill challenges. Thus, presenting them with the most appropriate and probable choices using the least number of necessary interactions is what this application achieves using machine learning algorithms. In a scenario where a caretaker wants to initiate a short conversation which can either be a question or a consultation, using trigger words the speech recognition algorithm will capture the English text and using NLP, predictive and most probable answers will be displayed on the screen which can be chosen by a click on the screen thereby reducing the effort of a motor-skill challenged person.

III. Word Prediction

As perhaps the highest profile application of NLP in

Augmentative and alternative communication, word prediction has maintained its place amongst all hearing and communication aids. In one early incarnation (Eulenberg et al., 1977), a frequency-tagged word list was used to suggest words for quick choices. This was in turn followed by more sophisticated systems using n-grams. Nearly all commercial AAC engines use some form of n-gram prediction. N-Gram Word Prediction reduces the number of selections required to generate a given text. Much of the cutting-edge word prediction research of the past few years has focused on sophisticated prediction engines capable of supporting multiple languages. Working on an English prediction system, Li and Hirst (2005) have selectively biased n-gram predicted words that are semantically related to recently appearing words. If word prediction is n-gram based, the approach where text input systems that exploit character probabilities is roughly analogous to using n-gram character prediction with large values of n. However, word-based character predictions are constantly opening up new possibilities for accurate character prediction.

A recent subject of research into word prediction has been the use of domain-specific predictions to boost keystroke savings, where the domain might be determined by our conversation, our geographical location or time of day. Trnka, Yarrington, McCoy, and Pennington (2006) have shown that combining topic-specific predictions with those derived from a general-purpose corpus can yield substantial gains in keystroke savings. Recent studies of genre-specific prediction have underscored the need to properly match the domain used for n-gram training to the type of text actually being generated by users (Wandmacher et al., 2008).

IV. Results and Discussions

HearMeOut is the only application in the market to have incorporated both gesture-based text prediction and pictogram-based augmentative and alternative communication using artificial intelligence to adapt to the needs of people with motor skill and verbal communication challenges. Sign Language Recognition Systems are developed using a methodology consisting of data acquisition and classification. HearMeOut uses a camera to

 DOI : 10.35291/2454-9150.2021.0031 © 2021, IJREAM All Rights Reserved.

capture live user motion which removes the needs of sensors in sensory gloves and thereby reduces the building cost too. The application incorporates Natural Language Conversation enabling the disabled users to engage in conversation using Speech Recognition. This feature comes in handy with the pictogram-based module which continuously listens to the environment looking for trigger words such as the name of the disabled person. Upon success, it activates communication. This application is built using a wide range of algorithms to afford flexibility and adaptability towards diverse motor and cognitive disabilities with utmost care towards data security.

V. ADVANTAGES OF HEARMEOUT

1. Users can set timeouts according to their comfort and play with the number of pictures displayed on screen. This is especially important for people with visual impairments, making images easier to be viewed.
2. High contrast images are used for pictogram-based communication which are important for people with cognitive impairments.
3. Predictive text is generated using state-of-the-art algorithms minimizing any chances of incorrect predictions.
4. Using AI to make context-based recommendations based on the user's location data and the time of day to predict what a person with a disability may wish to say in a specific context.
5. There is absolutely no user data storage minimizing data breaches except the pictograms which are accessed through a secured connection from the server.
6. A cost-effective and handy solution with a simple user interface enabling anyone to use the application.

VI. LIMITATIONS OF HEARMEOUT

1. A rich experience will involve the use of a large-scale gesture classification dataset with a large collection of images for the pictograms, thus leading to increased response time and effort.
2. Minimum connectivity is required to access the application else the predictions may take time to reflect resulting in loss of interest.

VII. CONCLUSION

Conventional sign language recognition systems have been developed from classifying only static signs and alphabets. However, with the help of artificial intelligence, HearMeOut can efficiently recognize dynamic movements that are fed to the system in the form of continuous sequences of images. Incorporating Natural Language Conversation with Gesture Based Language Generation, the proposed solution, HearMeOut, could play a crucial role in enabling people disabled with impaired hearing or speech or challenged with motor-skill movements with the chance to communicate with normal people using the LGenie language assistant with little concern and effort and once again integrate into the existing social structures without hesitation.

REFERENCES

[1] D. Jeffery Higginbotham PhD, Gregory W. Lesher PhD, Bryan J. Moulton MS & Brian Roark PhD (2012) The Application of Natural Language Processing to Augmentative and Alternative Communication, Assistive Technology, 24:1, 1424.

[2] Beukelman, D., Schley, H., Ternus, T., & Fager, S. (2010). New AAC access strategy: Gesture enhanced word prediction. American Speech, Language, Hearing Association, 2010 Convention, Philadelphia, PA.

[3] Goyal,Sakshi & Sharma,Ishita & Sharma, Shanu. Sign Language Recognition System for Deaf and Dumb People. International Journal of Engineering Research & Technology (IJERT). 2013 April; 2(4).

[4] Garofolo, John S., et al. CSR-I (WSJ0) Complete LDC93S6A. Philadelphia: Linguistic Data Consortium, 1993.

[5] Neamtu, Rodica & Camara, Andre & Pereira, Carlos & Ferreira, Rafael. (2019). Using Artificial Intelligence for Augmentative Alternative Communication for Children with Disabilities.

[6] Higginbotham, D. J., Lesher, G. W., & Luo, F. (2008). Using a web crawler to enhance AAC. *International Society for Augmentative and Alternative Communication*, Montreal, Canada. ISAAC.

[7] Hosom, J. P., Jakobs, T., Baker, A., & Fager, S. (2010). Automatic speech recognition for assistive writing in speech supplemented word prediction. In *Eleventh Annual Conference of the International Speech Communication Association*, Makuhari, Japan.

[8] Ichbiah, J. D. (1996). Method for designing an ergonomic one-finger keyboard and apparatus therefore. *U.S. Patent #No. 5,487,616*. Washington, DC: U.S. Patents and Trademarks Office.

[9] Kagohara, D.M., van der Meer, L., Achmadi, D., Green, V.A., O'Reilly, M.F., Mulloy, A., Lancioni, G.E., Lang, R., Sigafoos, J.: Behavioral intervention promotes successful use of an ipod-based communication device by an adolescent with autism. Clinical Case Studies 9(5), 328-338 (2010)

[10] Light, J., McNaughton, D., Beukelman, D., Fager, S.K., Fried-Oken, M., Jakobs, T., Jakobs, E.: Challenges and opportunities in augmentative and alternative communication: Research and technology development to enhance communication and participation for individuals with complex communication needs. Augmentative and Alternative Communication pp. 1-12 (2019)

[11] Pino, A., Kouroupetroglou, G.: Ithaca: An open-source framework for building component-based augmentative and alternative communication applications. ACM Transactions on Accessible Computing (TACCESS) 2(4), 14 (2010)

[12] Schlosser, R.: Roles of speech output in augmentative and alternative communication: Narrative review. Augmentative and Alternative Communication 19(1), 5-27 (2003)

[13] Schelhowe, H., Zare, S.: Intelligent mobile interaction: a learning system for mentally disabled people (imlis). In: International Conference on Universal Access in Human-Computer Interaction. pp. 412-421. Springer (2009)

 DOI : 10.35291/2454-9150.2021.0031 © 2021, IJREAM All Rights Reserved.

View publication stats