A Research Roadmap for Fair AI for People with Disabilities

K P SENTHILKUMAR M.E., C.S.E.,
KINGS ENGINEERING COLLEGE,
Opposite to Hyundai Motors,Chennai-Banglore Highway
Irungattukottai
Sriperumbudur 602117

ABSTRACT

People with disabilities (PWD) could be significantly impacted by AI technologies. Indeed, many cutting-edge AI systems, such as automatic speech recognition tools that can caption videos for people who are deaf and hard of hearing or language prediction algorithms that can improve communication for people with speech or cognitive disabilities, are motivated by the goal of improving the lives of PWD.Widely used AI systems, however, would not function well for PWD or, worse yet, might intentionally discriminate against them. These concerns about PWD fairness in AI haven't gotten much attention so far. We identify potential areas of concern in this position paper addressing how certain AI technology categories may affect specific disability communities if caution is not taken in their design, development, and testing. This risk analysis of the potential interactions between different types of AI and disabilities is meant to serve as a road map for future research that will be required to collect data, test these hypotheses, and develop more inclusive algorithms.

AuthorKeywords

Artificial intelligence, machine learning, data, accessibility, inclusiveness, justice and prejudice in AI, as well as moral AI.

INTRODUCTION

A significant difficulty is ensuring that AI systems function equitably for everyone as they become more and more prevalent in modern life. Unfair racial and gender bias in current AI systems has been found by researchers. Develop inclusive evaluation methods and practices, as well as recognize instances where homogeneous, exclusive data [9] or data reflecting negative historical biases is used for system training in order to better understand how AI systems function across diverse groups of individuals.

Even though many cutting-edge AI systems are motivated by the desire to improve the lives of people with disabilities (PWD)1 and even though these systems have the ability to alleviate many incapacitating conditions, concerns about AI's fairness for PWD have so far received little attention.

Fairness concerns for PWD may be more challenging to address than fairness issues for other groups, especially when the population's share of individuals with a given type of disability may be relatively modest.

They might be disregarded as outliers by existing AI algorithms even if they are included in training and evaluation data. Such problems could prevent PWD from using important technologies, unintentionally reinforce negative stereotypes about them (for example, if a chatbot learns to imitate someone with a disability), or even put their safety in danger (for example, if self-driving cars are not programmed to recognize pedestrians in wheelchairs).

To recognize and remedy weaknesses of AI systems for PWD, we suggest the research plan below: Create benchmark datasets to support replication and inclusion (and handle the complex ethical issues that creating such datasets for vulnerable groups might involve)Innovate new modeling, bias mitigation, and error measurement techniques and Identify ways in which inclusion issues for PWD may impact AI systems.

By considering how current important kinds of AI systems may require special consideration with respect to various classes of impairment, we move closer to the first of these goals in this position paper. The next critical step in developing AI that is inclusive of PWD is systematically examining the extent to which these interactions exist in practice, or demonstrating that they do not. However, identifying the scope of a problem is a prerequisite to solution.

Furthermore, we point out that it is a crucial question and one that may depend on the usage scenario of whether it is even ethical to develop specific categories of AI. Our discussion of various forms of AI is merely a description of how they might interact with disability, not a recommendation that such systems should be developed. In fact, there is a more extensive ethical debate to be had regarding how restricting certain forms of AI with negative connotations such as synthetic voices that could be used for deepfakes could disenfranchise PWD who could profit from such technology by restricting the ability to accurately mimic the voice of someone who is mute.

RISK ANALYSIS OF CURRENT AI SYSTEMS FOR DISABLED

Here, we categorize existing types of AI systems according to related features and name the demographics of people with disabilities who may find these systems to be troublesome. This risk assessment may not be comprehensive but is meant to serve as a springboard for further investigation. For instance, as new AI systems are developed, they would need to take handicap into account. Additionally, even though we made an effort to foresee potential reasons why certain disability groups might experience difficulties with certain classes of AI, we might not have fully identified all such groups. In fact, the "long tail" of disability and the possibility of co-occurring multiple disabilities are just two of the many reasons that making sure AI inclusion for PWD is particularly difficult.

Machine Learning

Computer vision systems examine images captured by fixed or moving cameras to spot patterns like the existence and characteristics of faces, bodies, or objects. When developing and evaluating the fairness of computer vision algorithms, it is vital to take into account disabilities that may have an impact on a person's outward appearance facial characteristics, facial expressions, body size or proportions, presence of assistive equipment, abnormal motion qualities, etc.

Face Identification

Face detection, identification, verification, and analysis are all possible with face recognition systems. Identification capabilities include spotting the presence of a face and/or drawing conclusions about its characteristics. Analysis capabilities include gender classification and emotion analysis. Face recognition technology is already in use in many different contexts, including biometric authentication security systems, criminal justice, interview support software and social/entertainment applications many of which are contentious.

If persons with different facial features and expressions are not taken into account while obtaining training data and assessing models, we believe that such strategies may not perform well for them. For those who have Down syndrome, achondroplasia, cleft lip/palate, or other diseases that cause recognizable facial variations, certain features of face analysis software may not be effective. Such systems may also malfunction for blind persons, which could lead to unexpected behaviors like a person not looking directly into a camera or different eye structure. It could also result in the person wearing dark glasses or other types of medical or aesthetic aids. That anticipated angle. Because people with autism or Williams syndrome may not express their emotions in a typical way, emotion processing algorithms may misinterpret their facial expressions. People with strokes, Parkinson's disease, Bell's Palsy, or other conditions that limit facial movement may also have difficulty interpreting facial expressions.

BodyRecognition

Body detection, identification, verification, and analysis are just a few of the capabilities that body recognition systems have access to in order to determine whether a body is present and/or draw conclusions about its characteristics. Body recognition systems can power applications using gait analysis e.g., for biometric authentication, sports biomechanics and path predictions used by self-driving cars, gesture recognition e.g., in VR and AR or gaming or both.

Body shape, posture, or movement variations in PWD may make it difficult for body recognition systems to function properly. For instance, gesture recognition systems are unlikely to function well for people with morphological differences e.g., a person with an amputated arm may be unable to perform bimanual gestures, may grip a device differently than expected, or their method of touching a screen may register an unexpected pattern. In situations when there is impairment, gesture recognition systems are also prone to fail. has an impact on how motion actually behaves, like in the case of someone who has tremors or spastic motion. Over time, fatigue may affect gesture performance and subsequently recognition accuracy, especially for populations that may be more vulnerable to it due to a disability or senior age. Differential gesture performance within or across days may also be caused by the scheduling of medications whose principal effects or side effects attenuate or enhance motor symptoms like tremor.

If body recognition is the only allowed interface, those who are completely immobile or who have significantly limited movement such as those with ALS or quadriplegia may be prevented from using certain technology. Additionally, body recognition systems may not be effective for individuals with mobility or morphology differences. For instance, if a self-driving car's pedestrian-detection algorithm does not include examples of individuals with posture variations caused by conditions like cerebral palsy, Parkinson's disease, advanced age, or who use wheelchairs during its training and evaluation, it may not correctly identify such individuals as objects to avoid or may incorrectly estimate the speed and trajectory.

Object,Scene,andTextRecognition

Optical character recognition (OCR) for objects, scenes, and Systems that can recognize text, handwriting, logos, and other everyday things provide labels, captions, and/or attributes (such location, activity, and relationship). Systems utilizing these capabilities have been widely accepted by PWD, especially persons with disabilities.

Since many of the photographs are acquired from social media platforms like Flickr, the majority of algorithms for identifying things from photos are trained using datasets of pictures taken by sighted individuals, and the pictures are frequently of good quality. Regarding geographic regions and household income, these picture data are known to be biased.

The error rates frequently rise when the models are used to process photos taken by a blind user since photos taken by blind individuals frequently have significantly lower quality than photos taken by sighted people due to issues with framing, blur, strange angles, poor lighting, etc.Similar issues can be seen in photographs shot by those who have tremors or other motor impairments. Furthermore, people with tremors or other motor impairments that affect writing neatness may have trouble using OCR models for handwriting recognition. The end-user experience of these tools may also not be adequately captured by the error metrics used to evaluate many vision systems, especially for end users with disabilities that might make it impossible for them to independently verify the system's output for example, a blind person would have to rely on an object detection system's output.

Speaking Devices

We use the term "speech systems" to refer to AI systems that can synthesize speech from symbolic inputs like text, Speech Synthesis Markup Language (SSML), or other encodings, or that can recognize the content (i.e., words) and/or features (i.e., prosody, speaker demographics) of speech. The accuracy and usefulness of speech systems may be hampered by conditions that affect a user's capacity to detect sound, as well as conditions that may affect the content or clarity of their speech.

Language Recognition

Systems that use automatic speech recognition ASR produce text from speech input. For those who are deaf or hard of hearing DHH, ASR systems have the potential to be crucial accessibility aids, such as by creating captions that can be overlaid as subtitles on movies or perhaps even by using augmented reality to live-caption face-to-face speech. For those who have trouble using their hands to operate conventional input devices, speech input is also helpful.

For those with abnormal speech patterns, ASR might not function properly. It is well known that ASR systems are biased; for instance, many of them work better for men than for women. Due to variances in pitch, pace, and clarity of speech by people of very advanced ages, many ASR systems today do not operate well for some older adults since they are not frequently represented in the training and evaluation of the systems.

Using current ASR techniques might be difficult for people with accents, particularly those caused by disabilities such as the "deaf accent while it is possible to train customized models for these groups. Dysarthria, another speech impairment, and the usage of speech-generating augmentative and alternative communication are all examples of speech impairments.

Speech Production

Text to speech (TTS) systems, which aim to produce realistic audio from symbolic inputs like text, SSML, or other markup, as well as cutting-edge AI tools like voice fonts [10, 48], which aim to accurately mimic the sound of a specific speaker, are examples of speech generation technologies. Many voice assistants, like Cortana, Alexa, Siri, and the Google Assistant, have TTS systems in place. TTS is also essential to many assistive technologies, such as screen readers for the blind and AAC devices for those with speech and mobility impairments. People with degenerative illnesses that cause progressive speaking loss may find value in voice banking to build unique voice fonts.

For example, individuals with cognitive or intellectual disabilities may need slower speech rates, whereas individuals with visual impairments may find rates too slow. System defaults for what constitutes comprehensible speaking rates may also need to be adjusted for specific disability segments. In the case of AAC technologies, text-based prediction techniques are frequently closely related to speech synthesis; therefore, the selection of training and evaluation corpora for prediction may need to be modified to be pertinent to the current needs and desired speech characteristics of AAC users, supporting expressivity and real self-representation.

Speaker Evaluation

Systems for speaker analysis can identify speakers, verify speakers, and draw conclusions about speakers' characteristics including age, gender, and emotion. There are several uses for speaker analysis systems, such as biometric authentication, improving speech transcription and personalization. Using visuals to promote sound awareness, speaker analysis systems have the potential to be crucial accessibility tools for people with DHH.

Tools for speaker recognition and speech analysis that draw conclusions about a user's personal attributes (such as gender or age) may not be effective for people with disabilities (PWD) whose speech sounds are considerably affected, such as dysarthria. For speakers with unusual prosody, such as those with autism or various forms of dementia, analysis algorithms that try to infer emotional state from prosodic traits are likely to fail.

Processing Text

Text analysis and translation are just two examples of the tasks that text processing systems carry out that are connected to understanding the content of text input. Text processing programs are probablysystems for minority languages used by disabled subcommunities, such as American Sign Language, pose accuracy and fairness difficulties for those with cognitive and or intellectual disabilities.

Text Evaluation

Text is fed into text analysis systems, which may look for content characteristics (such key phrases, named entities, and language) or author characteristics (including sentiment, personality, and demographics). Text analysis is widely used in pattern mining, information retrieval, and record management. Text analysis systems have the potential to be beneficial for people with disabilities (PWD) who have reading and writing difficulties, such as dyslexia, dysgraphia, or other cognitive differences, through visual illustration and focused highlighting or through intelligent spelling, grammar, and word or phrase suggestions.

The effectiveness and utility of many components of text analysis systems are likely to be impacted by cognitive and intellectual disability. For instance, there is some evidence to suggest that query rewriting and spelling correction software may not properly handle dyslexic spelling. Additionally, individuals with autism may express emotion differently in writing compared to neurotypical individuals, leading to inaccurate characterizations of their emotional state or personality. Text analysis systems can provide accuracy and fairness issues for individuals with cognitive and/or intellectual disabilities if these metrics are utilized as input to an automatic hiring system or automatic essay grading system, both of which are commonly employed with standardized aptitude testing.

AI integration

In addition to the aforementioned categories of systems for processing speech, text, and vision that concentrated on a single model, many complex AI systems include designs that integrate multiple models to provide more complicated behavior. Information retrieval and conversational agents are two prevalent instances of integrative AI that we will cover in this article.

Getting information back

In order to perform tasks like query rewriting, autocomplete suggestions, spelling corrections, search result ranking, content summarizing, and question answering, information retrieval (IR) technologies, such as those that run web search engines, rely on AI. IR systems' input and output might come in a variety of types, including as image, video, sound, or text.

Interactive Agents

End users can have conversations with conversational agents for a variety of real-world uses, such as customer service , education and health assistance. Various models, including as ASR, text analysis, TTS, and/or speaker analysis, are also used to power them. Conversational agents may ease users' workload when completing challenging activities and may offer cognitive support to those suffering from dementia or other intellectual disorders that impair memory or executive function.

Other AI Methods

It is important to evaluate risk factors for specific classes of AI applications, but it is also important to keep in mind that many AI techniques and practices that form the basis of such systems may result in biases against PWD. Examples include outlier detection techniques, practices for evaluating systems using aggregate metrics, definitions of objective functions, and the use of training data that does not accurately reflect the true complexity of the real world or the true use cases.

DISCUSSION

This position paper primarily focuses on the first stage of our research roadmap, which identified ways that (lack of) inclusion in training and evaluation of AI systems may adversely affect such systems' fairness for PWD. Our research roadmap included four proposed steps. To solve this, we talked about how common categories of AI could need to take different kinds of limitations into consideration.

CONCLUSION

The effectiveness and fairness of current classes of AI systems for individuals with disabilities may be constrained in this position paper by a number of strategies that serve as the foundation of AI. Our ultimate objective is the development of novel design principles, datasets, algorithmic approaches, and error measures that can assist AI systems in realizing their tremendous potential to assist PWD while avoiding the potential flaws we have described here. We expect that this study offers a research road map that will help practitioners and academics in AI develop systems that are efficient and fair to PWD.