HIGH SCHOOL EDITION
Journal of Student Research

Volume 13 Issue 2 (2024)

Exploring the Integration of Machine Learning and AI in the Treatment and Diagnosis of Communication Disorders

Luke Oh[1] and Gary Whitehead[#]

[1]Tenafly High School, USA
[#]Advisor

ABSTRACT

Communication disorders affect a significant portion of the population, and conventional diagnostic and treatment methods encounter challenges related to accessibility and efficiency. This paper examines how machine learning (ML) and artificial intelligence (AI) can be integrated into the diagnosis and treatment of communication disorders. The study emphasizes the potential transformation brought about by these technologies in improving patient outcomes, specifically highlighting successful initial diagnosis rates achieved with ML classifiers. Various novel ML methods developed over the years are discussed, offering the prospect of more precise treatment options for individuals with cognitive disabilities. However, the current research has limitations, including a narrow focus, the absence of a common gold standard, and difficulties in maintaining consistency across studies. Ethical considerations play a crucial role in the discussion, particularly concerning data privacy, security, and potential biases in ML models. The lack of diverse and representative datasets is noted as a factor that may lead to disparities in care. To advance the field, it is essential to replicate results across larger cohorts, address technical challenges, and improve the interpretability and transparency of AI-generated recommendations. These steps are identified as crucial for the ongoing development of ML and AI applications in the diagnosis and treatment of communication disorders.

Introduction

Communication disorders affect nearly five to ten percent of Americans and are prevalent in many lives (American Speech-Language-Hearing Association, n.d.). Almost twelve percent of U.S. children have communication disorders, with 34 percent having multiple communication disorders and 25.4 percent having other disorders in addition (Quick Statistics About Voice, Speech, Language, 2016). Communication disorder impairs a person's ability to receive, send, process, or comprehend ideas from verbal, nonverbal, or visual symbols. While the cause of a communication disorder may not be apparent, some common causes may be abnormal structures, neurological problems, oral-motor dysfunction, and many others (American Speech-Language-Hearing Association, n.d.-a).

Communication disorders are divided into four subtypes: speech disorder, language disorder, hearing disorder, and central auditory processing disorder. Speech disorders are disabilities of articulation with speech sounds, fluency, and voice; voice disorders are a common form of speech disorder distinguished by abnormal production of vocal quality, loudness, and resonance regarding an individual's age or sex. Language disorders impair understanding of spoken, written, or other symbol systems. Hearing disorders come from handicapped auditory sensitivity. Hearing disorders may restrict the comprehension, production, and preservation of speech. Hearing disorders are classified based on trouble detecting, recognizing, discerning, and grasping auditory information. Finally, central auditory processing disorders are shortfalls in correctly processing information from sound even though the person's hearing ability and intelligence are average. This disorder is not rooted in how well the individual hears sounds but in how their brain processes the imputed sounds. (American Speech-Language-Hearing Association, n.d.-a).

A complete physical exam, a psychometric logic test, or a speech test traditionally diagnose communication disorders (Giorgi, 2019). If treatment is needed, it is successful, with around 70% of the children benefiting (Lautieri, 2019). Treatment typically has three main goals: to help develop communication skills, build alternate strategies for when their communication abilities may be insufficient, and get recipients to practice and use their skills and techniques in different environments. Also, traditional communication disorder treatment falls into four primary treatment methods: speech therapy, behavior therapy, stimulant medication, and environmental modification (Lautieri, 2019). Speech therapy assists in building the recipient's vocabulary, organizing their thoughts, and correcting grammatical errors. Behavior therapy increases desired communication behavior and decreases unwanted issues and unhelpful coping methods. This therapy teaches valuable social skills through a system of rewards. Stimulant medication is an approach very similar to treating ADHD. Still, there is currently no drug that has been proven to consistently and effectively treat disorders, so only patients with high levels of irritability may receive some prescription (Posey et al., 2008). Environmental modifications are a crucial part of treatment as they support recipients and accommodate the setting to their needs to increase performance. For instance, children are greatly supported when their classroom environment is altered for their requirements, making them more successful learners and active participants (Lautieri, 2019).

Communication disorders, if untreated, can have very harmful effects. It can affect a person's quality of life, especially their mental well-being. Communication disorders have been recognized in several studies to be associated with depression, anxiety, substance use, and suicidal behavior. Botting et al. described (2016) high levels of depression and anxiety in 16-year-olds and 24-year-olds. Khurana et al. (2021) reveal a relationship between hearing disabilities and suicidal thoughts and attempts in 7546 residents over the age of 16 in Europe. Furthermore, a model adjusted for five "sociodemographic and clinical covariates" showed that hearing impairment nearly doubled the thought of suicide.

However, even though 1 in 12 US children are diagnosed with a communication disorder, nearly half of the US children do not receive any intervention (American Speech-Language-Hearing Association, n.d.-c), and many limitations cause this. The traditional speech-language evaluation for diagnosing communication disorders typically costs 150 to 400 dollars, and conventional treatment methods, such as therapy with a trained professional, may cost 65 to 175 dollars for every visit (Geller, 2023). Furthermore, cultural challenges, low trust in healthcare, and inadequate educational backgrounds contribute to the need for treatment. But, in addition to the difficulty of attaining treatment, Cunnings emphasizes a significant limit in effective therapy due to how patients have many issues during health visits. Health professionals like speech-language pathologists must tailor their communication style to the patient's unique situation. Still, there is limited research on people with communication disorders (Cummings, 2023).

Machine learning offers a viable solution to this prominent issue at present. Machine learning methods are automatic, effective, and consistent, allowing cheaper treatment plans. Unlike trained professionals who can be overbooked, which leads to increased costs and reduced accessibility, machine learning methods are not limited to such constraints.

Applications of Machine Learning and AI in Communication Disorders

Applications of Machine Learning to Diagnosing Communication Disorder

Although a trained professional is the most trusted to receive a diagnosis, they are challenging for many people. Machine learning and AI can be viable outlets for receiving an initial diagnosis, determining the severity of the condition, and assisting a professional in formulating a more precise determination.

Machine learning is a subclass of AI that self-learns without much human interference. When diagnosing communication disorders, Ml follows a general four-step structure: signal processing, feature extraction, classification, and results. First, in signal processing, an algorithm filters signals from a dataset, such as a voice recording, neural data, etc., to determine the main characteristics (Sharma et al., 2020). Next, in feature extraction, significant features are pointed in the signals to concise the data, making the classification process in the subsequent step work more efficiently. In classification, the chosen machine learning classifier analyzes the data from feature extraction and categorizes the data as "normal" or "disordered." Finally, these classification results can help pathologists produce more accurate diagnoses or identify the presence of communication disorders by themselves.

The application of machine learning has been tested on different data types and found a high success rate. Souissi et al. (Souissi & Chérif, 2016) tested their method on 2225 voice samples from the "Saarbrucken Voice Database." Their methods were based on two different classification models, the Support Vector Machine and Artificial Neural Networks. Both models were used after applying Mel Frequency Cepstral Coefficients, which performed feature extraction, and Linear Discriminant Analysis, which reduced the classifiers' complexity while improving the precision. In the end, the SVM and ANN models had an 86.44% and 87.82% accuracy rate in recognizing the presence of a communication disorder, proving that these two methods could efficiently identify the presence of voice disorders.

Ariyanti et al. (2021) also created a stacking ensemble method that builds on a signal binary classifier, such as SVM, because the scientists believe that binary classifiers are inferior when applied independently. Ensemble learning synthesizes several binary classifiers' results to produce a more accurate one. This solves the issue of having a binary, weak predictor by transforming many into one finer classifier. To create the ensemble model, the scientists require three elements: a base learner, which refers to the combination of binary classifiers (SVM); an algorithm that fits the data, meaning the algorithm confirms that the information from the base learner is compatible with the patterns by the data source; and a meta learner, another algorithm which retains the results from the base learner to create a final, composite prediction. The scientists stacked SVM models for their base model and used a deep neural network as their meta-learner. They tested their method on Fast Eastern Memorial Hospital patients, and the scientists' stacked ensemble approach accomplished an accuracy rate of 89.93%. This reveals the strength of a stacked approach rather than a binary one and also demonstrates the capacity as a method for voice classification.

Chet et al. investigated the brain morphological changes in patients with Parkinson's disease who eventually develop speech and voice disorders, referred to as hypokinetic dysarthria. Unlike using voice data like Souissi et al., Chet et al. utilized brain magnetic resonance imaging (MRI) scans from 134 Parkison patients. The scientists first applied feature selection and subsequently created a machine learning algorithm using a support vector training set — a training set is data used for the machine learning algorithm to identify rules and patterns for classification. The brain morphological changes were found to be associated with hypokinetic dysarthria, producing an exceptional prediction of the severity of hypokinetic dysarthria using machine learning-based neuroimaging.

Analyzing a variety of machine learning classifiers, Ibrahim et al. (2018) achieved high accuracy rates in identifying communication disorders in multiple countries. To acquire neural data, the scientists employed electroencephalography (EEG). EEG is a medical test used to measure the brain's electrical activity. Although the images are low in spatial resolution, meaning a worse ability to discriminate between different brain activities, they are inexpensive, widely accessible, and still an exceptional tool for analyzing brain activity.

Using the raw data from EEG images, the scientists preprocessed the data to filter out unwanted electrical signals that could interfere with accurate interpretations of the brain's activity. Then, the scientists took advantage of several techniques for feature extraction, as seen in Figure I, to locate the significant features of the EEG signals. Four different ML classifiers were applied to these filtered signals. After testing on different EEG datasets from Germany, the USA, and Saudi Arabia, the most successful technique was the combination of discrete wavelet transform plus Shannon Entropy, classified by the k-nearest neighbor (KNN) algorithm. This combination attained a 94.6% accuracy, illustrating that ML techniques have the potential to be applied to different data types in diverse countries.

To conclude this section, ML techniques are essential assets for diagnosing communication disorders. The techniques have been shown to have high accuracies applied for the two significant datasets, voice and neural, revealing the potential to create a diagnosis without a professional. This provides critical support for families who have to wait traditionally for years for a professional diagnosis.

Personalized Treatment Based on AI Data

As stated in the introduction, treatment is crucial and benefits around 70% of the people who take it. However, even seeing the benefits, only around half of the diagnosed patients receive treatment due to issues during health visits, difficulty in costs, and accessibility with planning health visits. ML and AI can be viable tools to unburden overwhelmed professionals as they can offer remote patient planning. They can also be integrated into treatments to create new techniques that improve the effectiveness of the treatment.

In speech therapy, automatic speech recognition (ASR) is used to treat dysarthria — a motor speech disorder that occurs from a neurological issue affecting the muscles used for speech (Website, 2023) — and help with articulation output. However, the standard ASR approaches do not provide enough reliable results, as they are highly vulnerable to speaker variability and challenging to obtain suitable and adequate speech data for non-English languages. Additionally, ASR trained on standard speech is not as helpful because its performance degrades significantly when the speech disorder worsens. So, Mulfari's deep learning-based approach called a convolution neural network (CNN), recognizes isolated words that show atypical speech using data from mobile software. This trained ASR system can be implemented in a tailored telerehabilitation, a treatment plan that allows patients to interact remotely with their pathologists to produce a highly personalized plan that reaps the benefits of Mulfari's CNN while supervised remotely by the speech-language pathologist. This greatly benefits the patient, as they receive an appropriate treatment that suits them the most, with CNN focusing on the small keywords that adjust the design as needed and also helps the pathologist, as their supervision does not require as much time as if they had to create an independent assessment without the CNN, and having to meet, in therapy sessions, constantly. (Mulfari et al., 2022)

Recently, three Syracuse researchers were granted 2.5 million dollars by the National Institutes of Health to address the "worldwide shortage of speech-language clinicians" by cultivating an automated system for treating speech sound disorders. Their system is rooted in the Speech Motor Chaining software, which accesses an extensive database of speech sounds and AI, enabling it to behave like a speech-language clinician. The researchers named their system "ChainingAI," the artificial intelligence form of the Speech Motor Chaining software. ChainingAI resolves the issues prevalent with other systems that detect speech errors. The dataset with audio for speech disorders, which Chaining AI works on, is much more extensive and has greater precision in recognizing the disorder. It can achieve its high accuracy because, from its database of 170,000 sounds, the AI trained itself by focusing on mispronounced "r" sounds and learning which patterns made irregular speech. During speech practice, the code of the software can then offer feedback through tongue positions and proper pronunciation. The system is not intended to replace speech clinicians. Still, it aims to significantly bolster the intensity of the treatment because, on top of the in-person sessions, the patients can work with the software and AI to have further practice, which the AI chooses based on what skills most need to be developed. (Stirling, 2023)

Emphasizing the burden that clinicians carry, Ghafghaz et al. outline that in order to create an effective treatment plan, the clinician must carefully observe the behavior of the patient, collect data, and distinguish problematic behaviors, which are all factors that can contribute to making errors and unreliable decisions. Another significant factor is that the patient may not be receptive to the treatment plan, especially for children.

This is why AI combined with technologies such as virtual reality to digitize the program is beneficial to children as it allows for further personalization.

Augmented reality, virtual reality, and mixed reality have been of recent interest because of their potential to aid patients by giving supplemental sensory details while also creating a controlled environment with an easy means to collect data. Additionally, studies illustrate that interactions in these virtual settings have boosted motivation. Especially with children being heavily connected to technology, AI implemented into phones can be used to help treat restrictive behaviors. For instance, children with behaviors that interfere with learning opportunities when playing video games can have their video games monitored by AI and redirected to another activity, which will prevent disruptive habits. Incorporating AI with these new technologies not only adds more options to individualize the treatment plan but also allows data-driven algorithms to be administered to improve data collection, further reducing the load on analysts as they may be able to save some time from collecting this data. (Ghafghazi et al., 2021)

Greater Monitoring using Artificial Intelligence

A part of having an effective treatment plan is adjusting it based on the patient's changing condition. Monitoring individuals who have or are suspected to have communication disorders involves a diligent process that requires observing, evaluating, and modifying. This time-consuming procedure may not be available to all since professionals are already overburdened with creating diagnoses and administering the treatment during practice sessions. This gap can be filled with AI, which can collect data and monitor patients continuously, utilizing sensors, apps, and other technologies. The AI can then adjust or have its data used by clinicians to tailor the plan to the patient's needs. Monitoring communication disorders using AI methods offers a promising approach to enhance healthcare delivery. By utilizing AI technologies like voice analysis, smartphone-based monitoring, and sensor data collection, healthcare professionals can continuously monitor patients, gather diverse data types, and adjust treatment plans effectively (Fitzpatrick et al., 2017). These tools allow for the analysis of voice features, physiological parameters, behavioral patterns, and environmental data to offer a comprehensive understanding of the patient's condition (Winzer et al., 2002). For example, AI has been successfully employed to identify speech patterns indicative of communication disorders in children, demonstrating its potential for early detection and intervention (Fiske et al., 2019).

AI has the potential to address the challenges healthcare professionals face in monitoring communication disorders, such as the time-consuming nature of traditional methods and the burden on clinicians (Jeong, 2020). By automating data collection and analysis processes, AI tools can alleviate the workload on healthcare providers, enabling them to concentrate on diagnosis and treatment (Petrovic & Maric, 2022). Moreover, AI algorithms can aid clinicians in tailoring treatment plans based on real-time data insights, resulting in more personalized and efficient care for patients (Urtnasan et al., 2021).

Case studies play a crucial role in showcasing the practical application of AI in monitoring communication disorders. For instance, research has illustrated that AI-driven voice analysis can detect speech sound disorders in individuals, underscoring the potential of AI-based automated speech therapy tools (Zhu & Pei, 2008). Additionally, integrating AI into speech therapy aims to streamline workflows, identify effective strategies, and implement best practices (Li et al., 2022). Collaboration across disciplines involving researchers, clinicians, and AI experts is essential for advancing the field of monitoring communication disorders. By combining expertise in healthcare, data science, and AI technologies, researchers can develop innovative solutions for effectively monitoring and managing communication disorders (Kong et al., 2016). Furthermore, research methodologies, such as utilizing machine learning algorithms to analyze voice data and predict speech disorders in individuals, play a pivotal role in advancing the field (Aylward et al., 2022).

Conclusion

In conclusion, the integration of machine learning and artificial intelligence in the diagnosis and treatment of communication disorders holds significant promise for transforming patient outcomes. The advancements discussed in this paper demonstrate the potential of these technologies to enhance diagnostic accuracy, personalize treatment plans, and offer continuous monitoring. However, several important considerations and limitations warrant attention as the field progresses. The significance of employing ML and AI technologies in the domain of communication disorders lies in their ability to revolutionize conventional diagnostic methods and treatment approaches. These innovations have the potential to improve the efficiency of diagnosis, enhance treatment personalization, and provide valuable insights for ongoing patient monitoring.

Despite the promising outcomes presented, it is essential to acknowledge the limitations inherent in the current state of research. The reliance on a limited number of test cases, often involving small populations, necessitates further investigation and validation across diverse cohorts. The absence of a common gold standard for assessment and the variability in defining speech naturalness across studies pose challenges to the generalizability and comparability of findings. The application of ML and AI in healthcare demands rigorous attention to ethical considerations, particularly concerning data privacy, security, and potential biases. Safeguarding personal and sensitive data is imperative to prevent unauthorized access or breaches. Moreover, efforts should be directed towards mitigating biases in ML models to avoid perpetuating disparities in the diagnosis and treatment of communication disorders.

To advance the field, future research should explore alternative sources, such as wearable technology, for continuous AI monitoring. Investigating precision in treatment options for individuals with cognitive disabilities is crucial for ensuring tailored interventions. Moreover, efforts should focus on replicating results across larger cohorts, addressing technical challenges, and enhancing interpretability and transparency of AI-generated recommendations. In summary, ongoing research and development are essential to overcoming current challenges, ensuring ethical use of AI in medical data, and promoting the integration of AI-assisted treatments into the standard practices of trained professionals.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

American Speech-Language-Hearing Association. (n.d.). Quick facts about ASHA. https://www.asha.org/about/press-room/quick-facts/#:~:text=5%25%20 to%2010%25%20 of%20 Americans,(source)

American Speech-Language-Hearing Association. (n.d.-a). Definitions of communication Disorders and Variations. https://www.asha.org/policy/rp1993-00208/#:~:text=I.,language%2C%20and%2For%20speech.

Ariyanti, W., Hussain, T., Wang, J.-C., Wang, C.-T., Fang, S.-H., & Tsao, Y. (2021). Ensemble and Multimodal Learning for Pathological Voice Classification. IEEE Sensors Letters, 5(7), 1–4. https://doi.org/10.1109/lsens.2021.3091141

Aylward, B. S., Abbas, H., Taraman, S., Salomon, C., Gal-Szabo, D. E., Kraft, C. A., … & Wall, D. P. (2022). An introduction to artificial intelligence in developmental and behavioral pediatrics.

Journal of Developmental &Amp; Behavioral Pediatrics, 44(2), e126-e134. https://doi.org/10.1097/dbp.0000000000001149

Bennett, T. D., Niedzwecki, C., Korgenski, E. K., & Bratton, S. L. (2013). Initiation of physical, occupational, and speech therapy in children with traumatic brain injury. Archives of Physical Medicine and Rehabilitation, 94(7), 1268-1276. https://doi.org/10.1016/j.apmr.2013.02.021

Botting, N., Toseeb, U., Pickles, A., Durkin, K., & Conti-Ramsden, G. (2016). Depression and Anxiety Change from Adolescence to Adulthood in Individuals with and without Language Impairment. PLOS ONE, 11(7), e0156678. https://doi.org/10.1371/journal.pone.0156678

Fiske, A., Henningsen, P., & Buyx, A. (2019). Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. Journal of Medical Internet Research, 21(5), e13216. https://doi.org/10.2196/13216

Fitzpatrick, K. K., Darcy, A. M., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. JMIR Mental Health, 4(2), e19. https://doi.org/10.2196/mental.7785

Geller, A. (2023, January 20). How much does speech therapy cost? Connected Speech Pathology. https://connectedspeechpathology.com/blog/how-much-does-speech-therapy-cost

Ghafghazi, S., Carnett, A., Neely, L., Das, A., & Rad, P. (2021). Ai-Augmented Behavior Analysis for children with developmental disabilities: Building toward precision treatment. IEEE Systems, Man, and Cybernetics Magazine, 7(4), 4–12. https://doi.org/10.1109/msmc.2021.3086989

Giorgi, A. (2019, November 5). Communication skills and disorders. Healthline. https://www.healthline.com/health/communication-skills-and-disorders#Symptoms

Ibrahim, S., Djemal, R., & Alsuwailem, A. (2018). Electroencephalography (EEG) signal processing for epilepsy and autism spectrum disorder diagnosis. Biocybernetics and Biomedical Engineering, 38(1), 16–26. https://doi.org/10.1016/j.bbe.2017.08.006

Khurana, M., Shoham, N., Cooper, C., & Pitman, A. (2021). Association between sensory impairment and suicidal ideation and attempt: a cross-sectional analysis of nationally representative English household data. BMJ Open, 11(2), e043179. https://doi.org/10.1136/bmjopen-2020-043179

Kong, G., Smith, P. H., Pilver, C. E., Hoff, R. A., & Potenza, M. N. (2016). Problem-gambling severity and psychiatric disorders among american-indian/alaska native adults. Journal of Psychiatric Research, 74, 55-62. https://doi.org/10.1016/j.jpsychires.2015.12.004

Lautieri, A. (2019, March 27). Treatment of communication disorders and recommended reading. MentalHelp.net. https://www.mentalhelp.net/disorders-of-childhood/treatment-of-communication-disorders-and-reading/

Li, H., Gao, J., & Feng, Y. (2022). Neuro-immune-endocrine mechanisms with poor adherence to aromatase inhibitor therapy in breast cancer. Frontiers in Oncology, 12. https://doi.org/10.3389/fonc.2022.1054086

Mulfari, D., La Placa, D., Rovito, C., Celesti, A., & Villari, M. (2022). Deep learning applications in telerehabilitation speech therapy scenarios. Computers in Biology and Medicine, 148, 105864. https://doi.org/10.1016/j.compbiomed.2022.105864

Petrović, S. and Maric, N. P. (2022). Improvement of the psychiatric care through outsourcing artificial intelligence technologies: where are we now?. Medicinska Istrazivanja, 55(2), 19-29. https://doi.org/10.5937/medi55-37718

Posey, D. J., Erickson, C. A., & McDougle, C. J. (2008). Developing drugs for core social and communication impairment in autism. Child and Adolescent Psychiatric Clinics of North America, 17(4), 787–801. https://doi.org/10.1016/j.chc.2008.06.010

Quick statistics about voice, speech, language. (2016, May 19). NIDCD. https://www.nidcd.nih.gov/health/statistics/quick-statistics-voice-speech-language#:~:text=Among%20children%20who%20have%20a,11%2D17%20have%20multiple%20disorders.

Sharma, G., Umapathy, K., & Krishnan, S. (2020). Trends in audio signal feature extraction methods. Applied Acoustics, 158, 107020. https://doi.org/10.1016/j.apacoust.2019.107020

Souissi, N., & Chérif, A. (2016). Artificial neural networks and support vector machines for voice disorders identification. International Journal of Advanced Computer Science and Applications, 7(5). https://doi.org/10.14569/ijacsa.2016.070546

Stirling, D. (2023, May 24). Researchers' artificial intelligence-based speech sound therapy software wins $2.5M NIH grant. SU News. https://news.syr.edu/blog/2023/05/24/researchers-artificial-intelligence-based-speech-sound-therapy-software-wins-2-5m-nih-grant/

Urtnasan, E., Joo, E. Y., & Lee, K. (2021). Ai-enabled algorithm for automatic classification of sleep disorders based on single-lead electrocardiogram. Diagnostics, 11(11), 2054. https://doi.org/10.3390/diagnostics11112054

Website, N. (2023, February 24). Dysarthria (difficulty speaking). nhs.uk. https://www.nhs.uk/conditions/dysarthria/

Winzer, K., Hardie, K. R., Burgess, N., Doherty, N., Kirke, D., Holden, M. T. G., … & Williams, P. (2002). Luxs: its role in central metabolism and the in vitro synthesis of 4-hydroxy-5-methyl-3(2h)-furanone. Microbiology, 148(4), 909-922. https://doi.org/10.1099/00221287-148-4-909

Zhu, J. and Pei, D. (2008). A luxp-based fluorescent sensor for bacterial autoinducer ii. ACS Chemical Biology, 3(2), 110-119. https://doi.org/10.1021/cb7002048