Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/41/20230738

An attempt at decrypting and understanding ASD language logic based on BERT with GPipe

Heming Chen[1], Chapman Kwong[2,4] and Xiang Li[3]

[1]School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Jiangsu, Suzhou, China
[2]Faculty of Information Technology, Cyber University, Tokyo, Japan
[3]School of Mechanical and Electrical Engineering, Shihezi University, XinJiang, Shihezi, China

[4]2201310097@cyber-u.ac.jp

Abstract. Autism spectrum disorder (ASD) has always been a mysterious cognition phenomenon that has yet to be solved and understood, as it is believed that ASD has a complex relationship with intelligence. Language can be seen as one way of manifesting intelligence. This paper wonders if decrypting and understanding ASD language logic gives a new answer to the ultimate question: Will Artificial Intelligence understand language? If they can, can they understand ASD language? If they cannot, would ASD language logic show a possible new path for AI to understand language? This paper uses BERT with GPipe, as BERT is well known for its language understanding abilities. This paper explores why there has not been any research in the field until now, as it later discovered there may be a phenomenon, neurotypical observer bias, that blinded views into knowing that thinking from an ASD perspective is something that exists. As test results, BERT is suspected to have possible neurotypical observer bias.

Keywords: Autism Spectrum Disorder, Language Logic, Neurotypical Observer Bias, Machine Learning, Natural Language Processing.

1. Introduction

For the past years, public views on autism spectrum disorder (ASD) have shifted from treating it as a disorder, a disease to a type of diversity. ASD is generally accepted as it has always been existing in human history. However, it was not until the early 1940s that ASD was first identified and described by Hans Asperger. Studies confirm that ASD is congenital and cannot be acquired later in life. It is also believed that ASD has a complex relationship with intelligence, as there are syndromes like the savant syndrome that, to this day, specialists cannot explain the causes. Modern analysis suggests that Alan Turing, the father of theoretical computer science and artificial intelligence, had strong autistic traits, and he would have been diagnosed with Asperger's (ASD) if he had lived today. Scientists such as Albert Einstein and Isaac Newton are also believed to have strong autistic traits and might have been diagnosed with ASD if they had lived today. It is undeniable that there is indeed a connection between ASD and intelligence, especially within the field of science.

In recent years, the development of artificial intelligence and its technological breakthroughs brought effective methods such as machine learning and natural language processing (NLP). Language can be

DOI: 10.54254/2755-2721/41/20230738

seen as a way of manifesting intelligence, as this is one of the main targets of NLP, asking the ultimate question of "Can artificial intelligence truly understand language." Neuroscience specialists followed this trend and started implementing NLP, hoping to understand ASD better. Unfortunately, current existing research and implementations of machine learning related to ASD topics mainly focus on "How to diagnose and identify ASD" using image classification and text translations instead of "actually trying to understand how ASD logic works" [1-6]. In other words, most research may have overlooked that the "ASD perspective" is worth discussing [7]. This is mainly due to two reasons. The first is that a phenomenon might cause it, Neurotypical Observer Bias. This means that most researchers and machine learning implementations towards ASD are from a neurotypical perspective, thus biased from the beginning. Even with NLP benchmarks like the General Language Understanding Evaluation (GLUE), it is still considered a relatively "NT perspective" standard and complete nonsense to ASD individuals [8]. The second reason is that typically speaking, ASD individuals suffer from human interaction and communication. Their social circles are naturally more minor, and they tend to be more cut off from each other and the world. Not having research on understanding ASD logic might be because there are insufficient datasets.

As this paper mentioned above, ASD is congenital, meaning that ASD individuals' language logic systems might be born differently. Thus, this paper started with this simple question: Will understanding ASD language logic be a new path toward understanding why ASD individuals think the way they do and eventually lead to a new way of improving artificial intelligence NLPs? If the father of artificial intelligence had a high possibility of being ASD, would thinking and observing the way he did lead to new opportunities?

This paper aims to give the public new scientific and humanistic perspectives towards ASD by creating an opportunity to open and broaden up the field of using machine learning, not only focusing on diagnosing and identifying but also trying to eliminate neurotypical observer bias and understand how ASD logic functions. This will discover hidden potential and provide ASD individuals with the needed environment by first analyzing and understanding their thoughts.

The datasets used in this paper are collected and self-made since no existing datasets cover this specific area. Datasets are collected across YouTube comments, Twitter tweets, blog posts, and Reddit posts. Samples that are related to ASD, such as ASD individuals sharing their own experiences with different language usage, different perspectives towards a particular topic, and different language understanding, are included. Samples related to neurotypical individuals, such as daily greetings involving "vague meanings," are included. No samples contain wordings directly pointing toward ASD individuals or communities.

This paper used BERT with GPipe to train the datasets as BERT has a more robust language understanding, hoping to see if BERT shows signs of neurotypical observer bias. This paper used GPipe to ensure time efficiency as the datasets will only get larger, and securing pipeline parallelism is crucial. This paper attempts to let people raise awareness of the potential of ASD while using ASD language logic to answer the ultimate question: Can NLP understand language? Will that day ever come? At the end of this paper, this paper shows that possible bias might exist, mainly subjecting to figurative language examples having a big influence when identifying the samples.

2. Methods

2.1. Data collection and processing

This paper attempts to create a fully ASD perspective dataset where "1" is ASD logic, meaning that these are what ASD individuals usually get as comments from others or how ASD logic thinks. The training dataset contains 50%-60% of "1" and mainly focuses on How ASD individuals use, interpret, and understand language. The test dataset contains random comments from various sources, with 20%-30% ASD-strong related data being "1." This paper aims to test whether BERT can distinguish the percentage of ASD-related comments without having any actual hints in the training dataset, such as words like ASD Asperger's, on the spectrum. The dataset is hand-picked by ASD individuals and

DOI: 10.54254/2755-2721/41/20230738

filtered by perspectives from both sides, non-ASD and ASD, as this would make the dataset as objective as possible. The dataset also went through several detailed revisions to better fit the needs, for example, limiting the number of figurative language samples as it is a well-known problem within ASD individuals and communities. Moreover, as mentioned above, this paper does not want data samples pointing directly to ASD individuals or communities.

2.2. Model description

This paper split a default BERT into three parts to fit GPipe and outputs to achieve pipeline parallelism. BERT is also well-known to be effective at text classification and sentiment analysis. Also, it is well-known that GPipe helps with BERT's performance as BERT needs to be as effective as possible with rather "vague" samples. As it is difficult to achieve effective results just with this paper, it focuses on achieving the fundamentals for further purposes and works. BERT, Bidirectional Encoder Representations from Transformers, is a natural language processing model developed by Google in 2018. Bidirectional means BERT processes text from left to right to left. BERT also uses the multi-head self-attention mechanism to learn complex relationships between tokens. GPipe is a pipeline parallelism library that can train giant neural networks effectively. GPipe evenly splits a network into multiple layers across multiple accelerators to achieve high and effective hardware utilization. It is well-known for its applications to multilingual translation.

3. Experimental results and analysis

Table 1. Default test dataset with mixed samples (Class 1 set at 90%).

	Accuracy	Precision	Recall	F score
Class 0	5.9%	100%	5.9%	11.1%
Class 1	100%	17.6%	100%	30.0%

Table 2. Class 1 contains figurative language samples and language logic samples (Class 1 at 80%).

	Accuracy	Precision	Recall	F score
Class 0	13.7%	94.2%	13.7%	24.0%
Class 1	95.8%	18.3%	95.8%	30.7%

Table 3. Only figurative language samples from default dataset are removed (Class 1 at 70%).

	Accuracy	Precision	Recall	F score
Class 0	29.7%	94.6%	29.7%	45.2%
Class 1	91.7%	20.8%	91.7%	33.9%

This paper used three different scales to test the performance. Table 1 is the default test dataset containing both figurative language and language logic samples, and the percentages of Class 1 and Class 0 are set at 90% to 10%. Tables 2 and 3 are two different variations from table 1. The result shows that accuracy drops when the proportion of Class 1 within the test dataset is decreased from 90% to 80%. This is expected to happen since the test dataset is intentionally set not to have any ASD-related wordings. However, train accuracy dropped significantly when only figurative language samples from the test dataset were removed and the percentage of Class 1 from 90% to 70%. This indicates that BERT might rely heavily on figurative language samples to identify Class 1 within the training dataset, which are ASD-related comments, instead of understanding the ASD language logic.

DOI: 10.54254/2755-2721/41/20230738

4. Discussion

This paper would like to expand the dataset into a more efficient size for further work. As mentioned above, results might need to be validated due to the small dataset size. In addition, figurative language samples might be effective in distinguishing but not as effective regarding segmentation and optimization. This paper wants to limit as many figurative language samples as possible, eventually eliminating it. The dataset needs more examples from an ASD perspective, and possible benchmarks and standards like The GLUE specifically for ASD need to be made. This paper has two critical limitations. First, the test dataset still needs to be ideal. Due to the severe lack of data, possibly too many figurative language examples are used. Since having difficulties with figurative language is a relatively well-known problem among ASD individuals, these data might have impacted the results, showing high accuracy [9, 10]. However, there are still some possible biases because the number of existing figurative language samples significantly impacts the results. Second, although this attempt suspects that BERT might have used specific words and hints that fit neurotypical observer bias to distinguish between 0 and 1 instead of understanding the language logic, the same problem with not having an ideal dataset exists.

5. Conclusions

In this paper, a new possible perspective towards the ASD community by showing that there is more to this community than just identifying and diagnosing, but also potential within because of how ASD individuals interpret the world and their differences with their underlying logic is discussed. This paper used BERT with GPipe to train an ASD logic dataset. Although the desired results might yet be obtained and this is more of an attempt, hope that through this attempt, people will increase awareness and gather data to expand on this field. This is just a beginning to broaden our horizons and ways of thinking, not just from a judgmental perspective. With enough awareness, more social values can be unearthed and used.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Eslami T, Almuqhim F, Raiker JS and Saeed F (2021) Machine Learning Methods for Diagnosing Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder Using Functional and Structural MRI: A Survey. Front. Neuroinform. 14:575999. doi: 10.3389/fninf.2020.575999

[2] Han Y, Rizzo DM, Hanley JP, Coderre EL, Prelock PA (2022) Identifying neuroanatomical and behavioral features for autism spectrum disorder diagnosis in children using machine learning. PLoS ONE 17(7): e0269773. https://doi.org/10.1371/journal.pone.0269773

[3] Kong X-J, Wei Z, Sun B, Tu Y, Huang Y, Cheng M, Yu S, Wilson G, Park J, Feng Z, Vangel M, Kong J and Wan G (2022) Different Eye Tracking Patterns in Autism Spectrum Disorder in Toddler and Preschool Children. Front. Psychiatry 13:899521. doi: 10.3389/fpsyt.2022.899521

[4] Hossain, M.D., Kabir, M.A., Anwar, A. et al. Detecting autism spectrum disorder using machine learning techniques. Health Inf Sci Syst 9, 17 (2021). https://doi.org/10.1007/s13755-021-00145-9

[5] Parlett-Pelleriti, C.M., Stevens, E., Dixon, D. et al. Applications of Unsupervised Machine Learning in Autism Spectrum Disorder Research: a Review. Rev J Autism Dev Disord (2022). https://doi.org/10.1007/s40489-021-00299-y

[6] Bone, D., Goodwin, M.S., Black, M.P. et al. Applying Machine Learning to Facilitate Autism Diagnostics: Pitfalls and Promises. J Autism Dev Disord 45, 1121–1136 (2015). https://doi.org/10.1007/s10803-014-2268-6

DOI: 10.54254/2755-2721/41/20230738

[7] Vulchanova M and Vulchanov V (2022) Rethinking Figurative Language in Autism: What Evidence Can We Use for Interventions? Front. Commun. 7:910850. doi: 10.3389/fcomm.2022.910850
[8] Tager-Flusberg, Helen & Rogers, Sally & Cooper, Judith & Landa, Rebecca & Lord, Catherine & Paul, Rhea & Rice, Mabel & Stoel-Gammon, Carol & Wetherby, Amy & Yoder, Paul. (2009). Defining Spoken Language Benchmarks and Selecting Measures of Expressive Language Development for Young Children With Autism Spectrum Disorders. Journal of speech, language, and hearing research : JSLHR. 52. 643-52. 10.1044/1092-4388(2009/08-0136).
[9] Kalandadze, T., Norbury, C., Nærland, T., & Næss, K.-A. B. (2018). Figurative language comprehension in individuals with autism spectrum disorder: A meta-analytic review. Autism, 22(2), 99–117. https://doi.org/10.1177/1362361316668652
[10] Kalandadze, T., Braeken, J., Brynskov, C. et al. Metaphor Comprehension in Individuals with Autism Spectrum Disorder: Core Language Skills Matter. J Autism Dev Disord 52, 316–326 (2022). https://doi.org/10.1007/s10803-021-04922-z