Personalized-ABA: Personalized Treatment Plan Generation for Applied Behavior Analysis using Natural Language Processing

Aman Kumar, Mareiko Au, Raj Semlawat, Malavica Sridhar, and Hitesh Gurnani
TheraDriver, Basil Therapy Inc.
{aman,mareiko,raj,mal,hitesh}@theradriver.com

Abstract

Autism Spectrum Disorder (ASD) is a neurological and developmental disability that affects how an individual learns, communicates, interacts with others. Applied Behavior Analysis (ABA) is a gold standard therapy for children and adults suffering from ASD to improve their learning, social, and communication skills. Today, 1 in 36 children are diagnosed with ASD with expectations that this rate will only continue to rise. The supply of certified ABA providers is alarmingly insufficient to meet the needs of children with ASD. In fact, waitlists to receive ABA therapy in the United States exceed 10 months in most states. Clinicians or Board Certified Behavior Analysts (BCBAs) are now experiencing intense bottlenecks around diagnostic evaluations and developing treatment plans quickly enough to support timely access to care. Over the past few years, Artificial Intelligence has changed the way industries operate by offering powerful ways to process, analyze, generate, and predict data. In this paper, we have addressed the problem of both time and supply restrictions faced by ABA providers by proposing a novel method for personalized treatment plan generation and program prediction by leveraging the capabilities of Deep Learning and Large Language Models (LLM). Additionally, we have introduced two separate models for behavior program prediction (F1-Score: 0.671) and skill acquisition program predictions (Rouge-1 Score: 0.476) which will help ABA providers in treatment plan implementation. Results are promising: an AI-generated treatment plan demonstrates a high similarity (Average Similarity Score: 0.915) to the original treatment plan written by a BCBA. Finally, as we partnered with a multi-state ABA provider in building this product, we ran a single-blind study that concluded that BCBAs prefer an AI-generated treatment plan 65 percent of the time compared to a BCBA-generated one.

1 Introduction

Over the past years, there has been a significant rise in the cases of Autism Spectrum Disorder (ASD). According to data collected by the Autism and Developmental Disabilities Monitoring (ADDM)(Maenner et al., 2023), 1 in 36 children in United States have autism. Increased awareness and screening, broadened diagnostic criteria, and better infrastructure for autism research have played a vital role in the rise of ASD prevalence. Applied Behavior Analysis (ABA) is regarded as a gold-standard therapy and is one of the most widely accepted therapies. Today, it is 100 percent covered by insurance and children are diagnosed as earlier as 2 years of age.

Effective ABA treatment relies on early diagnosis and effective treatment planning to individualize the behavior reduction goals and skill acquisition needs of every child. The type of ABA treatment plan is conventionally determined by a trained Board Certified Behavior Analyst (BCBA) via integrated assessment like VB-MAPP(CS et al., 2014), ABLLS(Partington and Analysts, 2010), Vineland, etc. and information derived from detailed patient intake forms, diagnostic reports, and the functional analysis of the patient. ABA treatment plans are comprehensive and tailored to meet the specific needs of individuals with Autism Spectrum Disorder (ASD). These plans encompass various components that work together to improve behaviors and enhance overall functioning. However, with a rise in ASD cases and increased demand for ABA services, there has been a shortage of ABA providers or Board Certified Behavior Analysts (BCBAs)(Chiri and Warfield, 2012)(Smith-Young et al., 2020). As the demand for ABA therapy escalates, maintaining a balance between workload and service quality therapy becomes increasingly complex for BCBAs. Due to this, individuals diagnosed with ASD face challenges in accessing the neces-

Proceedings of the 1st Workshop on NLP for Science (NLP4Science), pages 188–196
November 16, 2024

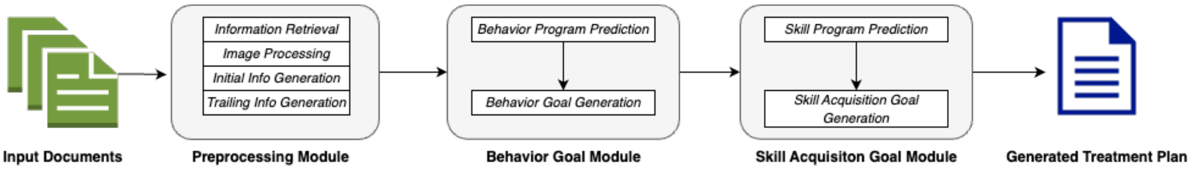

Figure 1: Illustration of the proposed method designed for end-to-end generation of treatment plan for ABA therapy personalized for each individual.

sary therapy and the required support. The scarcity of quality ABA therapy affects the progress of the individual in managing ASD, impacting mental health.

To address this problem, we have introduced a novel approach for creating individualized treatment plans tailored to the specific needs of any client using the advancements in the field of Deep Learning. We have also proposed two separate transformer-based(Vaswani et al., 2023) models to predict "Behavior Reduction" and "Skill Acquisition" programs, considering client's assessments (like VB-MAPP, ABLLS, etc.), diagnostic reports, and parent interviews. We have further shown our proposed models outperforms other state-of-the-art models in prediction tasks. The first-draft generated via our model will save a significant amount of time, as we are able to effectively analyze thousands of clinical documents to create and individualize a treatment plan. Through the experiments conducted for measuring the similarity of the AI-generated treatment plan versus the original one, the potential of the proposed method is revealed. The main contribution of the paper can be summarized as follows:

1. We propose an end-to-end novel method for generating a personalized treatment plan for an individual with ASD which takes Assessment Documents (like ABLLS, VB-MAPP), Parent Interview documents, and Diagnostic Reports as inputs. This will significantly reduce the time it takes a BCBA to create a treatment plan in addition to increasing the quality of goals for each child.

2. We also introduced two transformer-based models for the prediction of programs for the "Behavior Reduction" and "Skill Acquisition" Section.

3. Skill Acquisition program prediction uses an ensemble approach combining a rules-based and transformer-based model. We introduced our own word embedding model fine-tuned on the ABA treatment data for the ensemble approach using sentence-transformers(Reimers and Gurevych, 2019).

4. Treatment Plans generated using our approach show very high similarity with the original treatment plan and achieves an average similarity score of 0.915. When conducting a single-blind study with 35 BCBAs comparing AI-generated treatment plans vs. BCBA-generated treatment plans, BCBAs preferred AI-generated treatment plans 65 percent of the time. This strengthens our belief that AI can enable BCBAs to create the best version of a treatment plan for a child, as it is capable of processing and analyzing thousands of treatment plans in the past to create the most suitable and personalized treatment plan for a child. Furthermore, our models outperforms other state-of-the-art deep learning models used for similar tasks. To our knowledge, the proposed method for end-to-end treatment plan generation customized for an individual is first to be used in the domain of ASD.

2 Related Work

Artificial Intelligence and Deep Learning is increasingly used in the field of modern medicine for managing neurological conditions such as Alzheimer(EL-Geneedy et al., 2023)(Al Mamun et al., 2021) due to abundance availability of structured and unstructured data. The goal of Deep learning is to replicate cognitive abilities of human beings by analyzing complex datasets and generating meaningful pattern out of them without any human intervention(Pandey et al., 2022)(Egger et al., 2022). The healthcare sector has benefited from the advancements in the domain of deep learning with early diagnosis(Sorrentino et al., 2024), drug discovery(Carracedo-Reboredo et al., 2021), personalized treatment plan(Ng et al., 2021), etc. Recently, there has also been a used of deep learning for detecting depression(Fang et al., 2023) and managing mental health conditions(Shatte et al., 2019).

Recent advancements in the domain of deep learning has significantly impacted the field of Autism Spectrum Disorder (ASD) by offering new means for diagnosis and treatment. (Kollias et al., 2022) proposes a method for ASD detection using eye movement of an individual. Deep learning, particularly involving neural networks have shown promising results in the field of ASD. For instance, (Ahammed et al., 2021) uses Convolutional Neural Network (CNN) for classification of ASD on functional MRI data. Similarly, there has also been use of Recurrent Neural Networks (RNN)(Sudha and Vijaya, 2021) and Long Short-Term Memory (LSTM) for analyzing time-series data for diagnosis of ASD. For instance, (Li et al., 2019) uses LSTM to diagnose children with ASD based on raw video data. Deep learning techniques are also used for dimensionality reduction and feature extraction which helps in ASD research. (Kim et al., 2021) uses Variational auto-encoder (VAE) for representation learning, enhancing the interpretability of complex datasets. Personalization of treatment(Kohli et al., 2022) for an individual with ASD has also benefited from the advancements in deep learning while there has also been work on Ensemble learning for classification of ASD(Gaur et al., 2023). As the field of deep learning continue to evolve, it can play a crucial role in ASD research.

3 Methodology

This section presents the proposed method for an end-to-end treatment plan generation personalized for each individual (as depicted in Figure 1) along with the details about the prediction of behavior and skill acquisition programs. It takes the raw assessment file (generally in excel format for ABLLs, VB-MAPP, etc), diagnostic reports, and the parent interview documents of the client as inputs. Our proposed method has been divided into different sections for pre-processing and the analysis of the input, prediction, and generation of goals tailored to meet the specific requirements of the client.

3.1 Preprocessing of Input

We take raw assessment files, diagnostic reports and parent interview forms as inputs for our proposed method. We then extract the following information from the corresponding document for formulation of the treatment plan:

- Client Demographics and Diagnostic Code: Diagnostic Report

- Client Medical History, Language or Communication Skills, Social or Play Skills, Repetitive, Rigid, Restrictive or Challenging Behavior: Parent Interview Docs

We use image processing and with the help of opencv(Bradski, 2000) and openpyxl(Clark and Gazoni, 2010), we analyze the raw assessment files and convert them into a summary table containing domains(like Mand, Tact, Intraverbal, etc.) and score of the client in that particular domain. This saves a significant amount of time for BCBAs, who typically spend time analyzing these assessments and converting them into structured formats manually. In fact, we surveyed roughly 40 BCBAs who quantified that it takes between 10-15 hours to create a treatment plan for a child in their caseload. We believe our AI-generated treatment plan will bring this time down to 1 hour or less.

3.2 Behavior Program Prediction

In ABA therapy, Behavior Programs are focused on addressing specific behavior of an individual which maybe interfering with their daily life routine. These goals are aimed to decrease the challenging behavior and increase the desired behavior.

We formulated the problem of Behavior Goal Prediction as a multi-label classification tasks which takes Language/Communication Skills, Social/Play Skills and Restrictive or Challenging Behavior of the client as an input and outputs the set of the behavior program which interferes with client's daily life and functioning and which needs to be addressed in the ABA therapy. We used a transformer based model named Deberta-v3(He et al., 2021) for the multi-label classification problem. Deberta-v3 is an improvement over BERT(Devlin et al., 2019) and RoBERTa(Liu et al., 2019) based models by using disentangled attention and enhanced mask decoder. We train the model on our curated dataset using binary cross entropy loss which measures the dissimilarity between the true labels and the predicted probabilities (Eq.1).

$$Loss = (Y)(-log(Y_p)) + (1\text{-}Y)(-log(1\text{-}Y)_p) \quad (1)$$

where Y_p = predicted probability of the class and Y = true label.
We calculate the loss for each class and then sum over all the classes for training our model.

3.3 Skill Acquisition Program Prediction

In ABA Therapy, Skill Acquisition Goals play a crucial role by providing a structured framework, which helps the client with developmental or behavioral challenges and helps them gain essential skills including communication, social skills, and self-management.

3.3.1 Transformer-based Model

Since the number of skill acquisition programs can be very large in number, it was not feasible to train a classification model for this task with a limited dataset. Hence, we formulated the problem as a question-answering task, where the input is the client information which includes their demographics information, previous medical history, clinical and home observation data and assessment results (like VB-MAPP, ABLLs, Vineland, etc.) and output is the set of skill acquisition programs suited for the particular client. We used FlanT5(Chung et al., 2022) which is an encoder-decoder(Vaswani et al., 2023) based model for this question answering task using a labeled dataset. Here, our labeled dataset consisted of a question **x**, and a response **y**(list of skill acquisition programs) corresponding to the given question or client information. Consequently, the training process is aimed at minimizing the cross-entropy loss between the predicted probabilities and true class labels(Eq.2.)

$$Loss(\theta) = -\,E[log p_\theta(y|x)] \tag{2}$$

3.3.2 Rule-based Model

The rules-based model for predicting skill acquisition programs leverages assessment summary tables created from raw assessment files for generating tailored recommendations for skill development of an individual. The model employs a structured set of rules derived from gold-standard assessment curriculum guides, then validated with the help of BCBA domain expertise. The model identifies the client-specific deficits in each domain by analyzing detailed data from raw assessment files and then translate these identified deficits into set of skill acquisition goals using the set of predefined rules. This approach ensures that goal predictions are tailored to unique needs of an individual.

3.3.3 Ensemble Model

The Ensemble Model for predicting skill acquisition programs integrates both the transformer-based and rules-based model. We create a client vector which includes the demographic information of the client and their home and clinical observation results. We created our own embedding model by fine-tuning the ClinicalBERT(Alsentzer et al., 2019) model to improve its performance on our specified use case. Our dataset consists of a pair of sample in the format $\{x1, x2, s\}$, where *x1* and *x2* are pair of sentences and *s* is a binary label. *s* = 1 if *x1* and *x2* are deemed similar, and *s* = 0 if they are deemed dissimilar. We utilize sentence-transformers for fine-tuning and use contrastive loss between the input vector consisting of client demographic information, home and clinic observation results(*x1*) and the program list(*x2*). The client vector was created from our fine-tuned embedding model. We then calculate the similarity of each of the predicted skill acquisition program from both the approaches (Transformer-based and rule-based model) with the client vector. Final set of results contains those programs which posses similarity score of greater than 0.5 with the client vector.

3.4 Goals Generation

In our proposed method, we utilized GPT-3.5 Turbo(Liu et al., 2023) for generating the short-term and long-term goals corresponding to each of the program. In our research, we implemented Retrieval-Augmented Generation (RAG) combined with Prompt Engineering with GPT-3.5 Turbo for enhancing the quality of the generated goals. We developed a specialized custom database tailored to our specific need which includes client information and short-term and long-term goals corresponding to that client. Retrieval-Augmented Generation (RAG) enabled the GPT model to utilize the custom-built database of relevant information to inform and enhance the goal generation process. Prompt engineering involves designing and refining input prompts to guide the model in producing more relevant and accurate outputs. This process included providing explicit instructions to the model, experimenting with different prompt structures, and iteratively refining prompts based on the model's responses. Combining RAG and Prompt Engineering, we aimed to leverage GPT's advanced language capabilities for generating more accurate and precise goals specific to a particular

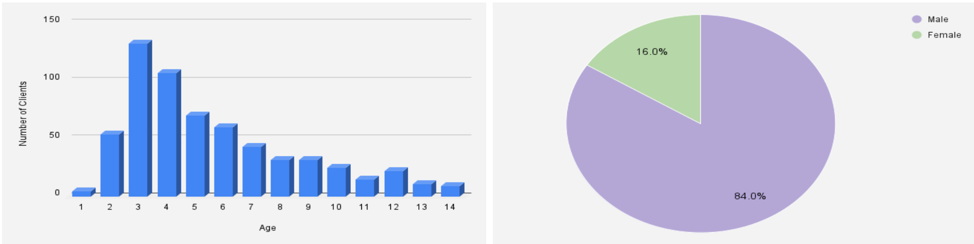

Figure 2: The left figure presents the number of client in each age group, with bars indicating the count of clients. Right figure presents a pie chart indicates the percentage distribution of male and female clients relative to the total client population.

individual.

4 Experiments

4.1 Experimental Settings

4.1.1 Models and Datasets

We evaluate the effectiveness of our proposed method using the following models for program prediction task:

- Behavior Program Prediction: Deberta-v3-base. The DeBERTa V3 base model has 12 layers and 86M backbone parameters with a vocabulary size of 128k tokens and is trained on 160GB of data as Deberta v2.

- Skill Program Prediction: Flan-t5-base. Flan-T5 is an improvement over T5-based models. FLan-T5-base contains 12 hidden layers and 250M model parameters.

We further calculate the semantic textual similarity of the final treatment plan generated with the original treatment plan using pre-trained sentence transformer models.

We curated a custom dataset for validating the effectiveness of our proposed method. All the clients included in the study have been diagnosed with Autism Spectrum Disorder (ASD) and are within the age range of 2 to 14 years. The full dataset encompasses a total of 617 clients which were further divided into training and testing set with an 80:20 split. The distribution of clients by age group and gender proportions is shown in Figure 2. For each client, the dataset includes the original treatment plan formulated by BCBA and other input documents like the parent interview form, diagnostic report, and raw files of the assessment which were conducted for that particular client, for example: VB-MAPP, ABLLS, Vineland, etc.

4.1.2 Implementation Details

We trained our model for behavior programs and skill program predictions using 4 Nvidia T4 GPUs using the Pytorch deep learning framework. We used the Accelerate library to make our distributed training easier, more effective, and efficient at the same time. Batch size has been set to 4 for all the experiments. The learning rate is set to 2e-5 for behavior program prediction , and 3e-4 for skill acquisition program prediction using a cosine rate scheduler. Pre-processing steps including information retrieval from input documents and analysis of raw assessment files were carried out in CPU(Apple M3 Pro chip: 12-core CPU).

4.1.3 Evaluation Setting

All the evaluation were carried out on the test set of our curated dataset. For the quantitative performance comparison, we adopt an F1-score, Rouge score and Exact Match as an evaluation metric. The F1-score is specifically developed for assessing the performance of a classification model, while Rouge score are used to evaluate the quality of machine-generated text. A higher F1 and Rouge score indicated better model performance. Exact Match (EM) is another question-answering evaluation metric that only gives two scores (0 or 1). EM score is 1 if the generated answer is precisely the same as the predicted answer, else, it gives 0. We also evaluate the effectiveness of our end-to-end treatment plan generation approach by calculating its similarity with the original treatment plan. We used sentence transformers for this approach which outputs a similarity score in range of -1 to 1, where -1 indicates complete dissimilarity while 1 indicates complete similarity.

Model	Params	Rouge-1	Rouge-2	Rouge-l	Exact Match
Flan-T5-base(Chung et al., 2022)	250M	0.0139	0.0	0.010789	0.0
Flan-T5-large(Chung et al., 2022)	780M	0.01986	0.0	0.01708	0.0
Phi-2b(Abdin et al., 2023)	2.7B	0.05830	0.0043	0.03389	0.0
Ours(based on Flan-T5-base)	**250M**	**0.4762**	**0.3583**	**0.3764**	**0.3376**

Table 1: Performance comparison for Skill Program Prediction using different models. It can be observed that our model outperforms all the other models on rouge score and exact match metric having least number of parameters compared to other models.

Model	Params	F1-score
Deberta-v3-base(He et al., 2021)	86M	0.1023
Deberta-v3-large(He et al., 2021)	304M	0.0662
Roberta-base(Liu et al., 2019)	125M	0.087
Biobert v1.1(Lee et al., 2019)	110M	0.19535
Ours(Based on Deberta-v3-base)	**86M**	**0.671**

Table 2: Table shows the F1-score comparison of different model for behavior program prediction. Our model achieves better results compared to other models with minimal number of parameters.

Name	Minilm-L6-v2	Bert-base	Mpnet-base
Client 1	0.883945	0.97367	0.950932
Client 2	0.84949	0.92193	0.923908
Client 3	0.949224	0.97867	0.95126
Client 4	0.93826	0.962447	0.95891
Client 5	0.90818	0.939519	0.94793
Client 6	0.746431	0.82151	0.858006
Average	**0.879255**	**0.93296**	**0.93182**

Table 3: Table shows the semantic textual similarity score between the treatment plan generated using our approach and the original treatment plan for 6 random clients from the test set. Similarity score was calculated using sentence-transformers using different models. It can observed that our treatment plan shows very high similarity score with the original treatment plan hence validating the effectiveness of our proposed method.

4.2 Experimental Results

4.2.1 Behavior Program Prediction

We evaluate our model trained for behavior program prediction using Microsoft Deberta-v3-base model on the test set of our custom curated dataset. We compare its performance with similar other models as shown in Table 2. We observe that our approach is able to beat all other models by a huge margin and has fewest number of parameters, hence making it a suitable choice for deployment in a resource-constrained environment.

4.2.2 Skill Acquisition Program Prediction

We evaluate our model trained for skill acquisition program prediction using Google Flat-T5-base model on the test set of our custom curated dataset. Table 1 shows the performance of our model using Rouge Score and Exact Match evaluation metric. We observe that our model is able to outperform all the other models which validates the effectiveness of our approach.

4.2.3 Full Treatment Plan Generation

We further evaluate the effectiveness of our proposed method for entire treatment plan generation by measuring the similarity between generated and original treatment plan as show in Table. We took 6 random client from the test set and generated the full treatment plan by taking diagnostic report, parent interview and raw assessment files as an input. We then calculated the similarity using different models with the original treatment plan for better validation using sentence-transformers. We observe that the treatment plan generated using our proposed method posses very high similarity with the original treatment plan, which validates the effectiveness of our proposed methodology. When sharing AI-generated treatment plans versus BCBA-generated treatment plans via a single blind study to 35 BCBAs, our AI-generated treatment plan was preferred 65 percent of the time.

4.3 Ablation Studies

Our proposed method for an end-to-end treatment plan generation shows promising results and also outperforms state-of-the-art NLP models on behavior program and skill acquisition program prediction tasks. We performed extensive ablation experiments to demonstrate the effectiveness of the proposed method on personalized downstream tasks. We tried an exhaustive hyperparameter search and fine-tuning the learning rate. All these modifications results in a very negligible change in the overall performance of the model. We also tried modifying the loss function by incorporating ABA Therapy rules, however, no significant changes were observed in model performance.

5 Conclusion

In this work, we propose a novel method for full treatment plan generation specific to the needs of a particular individual. Our proposed method outperforms other state-of-the-art model in similar domain in program prediction tasks. We further validate the effectiveness of our approach by calculating the similarity of the generated treatment plan with the original treatment plan. The treatment plan generated using our approach possess very high similarity with the original treatment plan. Our proposed method will not only help in automating the treatment plan generation process but will also reduce the time taken by a BCBA in formulating the treatment plan. On average, a BCBA spends anywhere between 10-15 hours creating a single treatment plan from start to finish. We believe that our AI-generated treatment plan will enable them to do so in 1 hour or less. The applicability of our proposed method on other neurodevelopmental disorders outside of autism are open avenues for future work.

6 Limitations

While our proposed end-to-end method for treatment plan generation achieves a very high semantic similarity score with the original treatment plan and our proposed models for behavior program prediction and skill acquisition program prediction generates better results, it does have certain limitations that we plan to address in future. Since our curated dataset contains client suffering from ASD, the applicability of our approach to similar neurological problems is still unexplored. We used deberta-v3 as a backbone for multi-label classification task in the domain of ASD. However, validating the effectiveness of techniques like label clustering(Ding et al., 2020) and graph-label attention(Pal et al., 2020) in the domain of ASD can be the scope of future work. There is also a lack of labelled dataset due to scarcity of available open-source data in the domain of ASD. Furthermore, the variances in the response of BCBAs or ABA providers can be one of the areas of research in the future for improving the robustness of the model. To address this, we plan to leverage reinforcement learning techniques to address and incorporate the individualized preferences of BCBAs when generating a treatment plan for a child in their caseload.

References

Marah Abdin, Jyoti Aneja, Sebastien Bubeck, Caio César Teodoro Mendes, Weizhu Chen, Allie Del Giorno, Ronen Eldan, Sivakanth Gopi, Suriya Gunasekar, Mojan Javaheripi, Piero Kauffmann, Yin Tat Lee, Yuanzhi Li, Anh Nguyen, Gustavo de Rosa, Olli Saarikivi, Adil Salim, Shital Shah, Michael Santacroce, Harkirat Singh Behl, Adam Taumann Kalai, Xin Wang, Rachel Ward, Philipp Witte, Cyril Zhang, and Yi Zhang. 2023. Phi-2: The surprising power of small language modelss.

Md Shale Ahammed, Sijie Niu, Md Rishad Ahmed, Jiwen Dong, Xizhan Gao, and Yuehui Chen. 2021. Darkasdnet: Classification of asd on functional mri using deep neural network. *Frontiers in Neuroinformatics*, 15.

Shamim Al Mamun, M. Shamim Kaiser, and Mufti Mahmud. 2021. An artificial intelligence based approach towards inclusive healthcare provisioning in society 5.0: A perspective on brain disorder. In *Brain Informatics: 14th International Conference, BI 2021, Virtual Event, September 17–19, 2021, Proceedings*, page 157–169, Berlin, Heidelberg. Springer-Verlag.

Emily Alsentzer, John R. Murphy, Willie Boag, Wei-Hung Weng, Di Jin, Tristan Naumann, and Matthew B. A. McDermott. 2019. Publicly available clinical bert embeddings.

G. Bradski. 2000. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Paula Carracedo-Reboredo, Jose Liñares-Blanco, Nereida Rodríguez-Fernández, Francisco Cedrón, Francisco J. Novoa, Adrian Carballal, Victor Maojo, Alejandro Pazos, and Carlos Fernandez-Lozano. 2021. A review on machine learning approaches and trends in drug discovery. *Computational and Structural Biotechnology Journal*, 19:4538–4558.

Giuseppina Chiri and Marji Erickson Warfield. 2012. Unmet need and problems accessing core health care services for children with autism spectrum disorder.

Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Eric Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. 2022. Scaling instruction-finetuned language models.

Charlie Clark and Eric Gazoni. 2010. Openpyxl: A python library to read/write excel 2010 xlsx/xlsm files. Computer software.

Barnes CS, Mellor JR, and Rehfeldt RA. 2014. Implementing the verbal behavior milestones assessment and placement program (vb-mapp): Teaching assessment techniques.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding.

Fei Ding, Xin Kang, Shun Nishide, Zhijin Guan, and Fuji Ren. 2020. A fusion model for multi-label emotion classification based on BERT and topic clustering. In *International Symposium on Artificial Intelligence and Robotics 2020*, volume 11574 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 115740D.

Jan Egger, Christina Gsaxner, Antonio Pepe, Kelsey L. Pomykala, Frederic Jonske, Manuel Kurz, Jianning Li, and Jens Kleesiek. 2022. Medical deep learning—a systematic meta-review. *Computer Methods and Programs in Biomedicine*, 221:106874.

Marwa EL-Geneedy, Hossam El-Din Moustafa, Fahmi Khalifa, Hatem Khater, and Eman AbdElhalim. 2023. An mri-based deep learning approach for accurate detection of alzheimer's disease. *Alexandria Engineering Journal*, 63:211–221.

Yuqi Fang, Mingliang Wang, Guy G. Potter, and Mingxia Liu. 2023. Unsupervised cross-domain functional mri adaptation for automated major depressive disorder identification. *Medical Image Analysis*, 84:102707.

Manu Gaur, Kunal Chaturvedi, Dinesh Kumar Vishwakarma, Savitha Ramasamy, and Mukesh Prasad. 2023. Self-supervised ensembled learning for autism spectrum classification. *Research in Autism Spectrum Disorders*, 107:102223.

Pengcheng He, Jianfeng Gao, and Weizhu Chen. 2021. Debertav3: Improving deberta using electra-style pretraining with gradient-disentangled embedding sharing.

Jung-Hoon Kim, Yizhen Zhang, Kuan Han, Zheyu Wen, Minkyu Choi, and Zhongming Liu. 2021. Representation learning of resting state fmri with variational autoencoder. *NeuroImage*, 241:118423.

Manu Kohli, Arpan Kar, Anjali Bangalore, and A P Prathosh. 2022. Machine learning-based aba treatment recommendation and personalization for autism spectrum disorder: an exploratory study. *Brain Informatics*, 9.

Konstantinos-Filippos Kollias, Christine K. Syriopoulou-Delli, Panagiotis Sarigiannidis, and George F. Fragulis. 2022. Autism detection in high-functioning adults with the application of eye-tracking technology and machine learning. In *2022 11th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, pages 1–4.

Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang. 2019. Biobert: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4):1234–1240.

Jing Li, Yihao Zhong, Junxia Han, Gaoxiang Ouyang, Xiaoli Li, and Honghai Liu. 2019. Classifying asd children with lstm based on raw videos. *Neurocomputing*, 390.

Yiheng Liu, Tianle Han, Siyuan Ma, Jiayue Zhang, Yuanyuan Yang, Jiaming Tian, Hao He, Antong Li, Mengshen He, Zhengliang Liu, Zihao Wu, Lin Zhao, Dajiang Zhu, Xiang Li, Ning Qiang, Dingang Shen, Tianming Liu, and Bao Ge. 2023. Summary of chatgpt-related research and perspective towards the future of large language models. *Meta-Radiology*, 1(2):100017.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized BERT pretraining approach. *CoRR*, abs/1907.11692.

Matthew J Maenner, Zachary Warren, Ashley Robinson Williams, Esther Amoakohene, Amanda V Bakian, Deborah A Bilder, Maureen S Durkin, Robert T Fitzgerald, Sarah M Furnier, Michelle M Hughes, Christine M Ladd-Acosta, Dedria McArthur, Elise T Pas, Angelica Salinas, Alison Vehorn, Susan Williams, Amy Esler, Andrea Grzybowski, Jennifer Hall-Lande, Ruby H N Nguyen, Karen Pierce, Walter Zahorodny, Allison Hudson, Libby Hallas, Kristen Clancy Mancilla, Mary Patrick, Josephine Shenouda, Kate Sidwell, Monica DiRienzo, Johanna Gutierrez, Margaret H Spivey, Maya Lopez, Sydney Pettygrove, Yvette D Schwenk, Anita Washington, and Kelly A Shaw. 2023. Prevalence and characteristics of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, united states, 2020.

Kenney Ng, Uri Kartoun, Harry Stavropoulos, John A Zambrano, and Paul C Tang. 2021. "personalized treatment options for chronic diseases using precision cohort analytics.

Ankit Pal, Muru Selvakumar, and Malaikannan Sankarasubbu. 2020. Magnet: Multi-label text classification using attention-based graph neural network. In *Proceedings of the 12th International Conference on Agents and Artificial Intelligence*. SCITEPRESS - Science and Technology Publications.

Babita Pandey, Devendra Kumar Pandey, Brijendra Pratap Mishra, and Wasiur Rhmann. 2022. A comprehensive survey of deep learning in the field of medical imaging and medical natural language processing: Challenges and research directions. *Journal of King Saud University - Computer and Information Sciences*, 34(8, Part A):5083–5099.

J.W. Partington and Partington Behavior Analysts. 2010. *The Assessment of Basic Language and Learning Skills-revised (the ABLLS-R)*. The Assessment of Basic Language and Learning Skills-revised. Behavior Analysts.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Adrian B. R. Shatte, Delyse M. Hutchinson, and Samantha J. Teague. 2019. Machine learning in mental health: a scoping review of methods and applications. *Psychological Medicine*, 49(9):1426–1448.

Joanne Smith-Young, Roger Chafe, , and Rick Audas. 2020. "managing the wait": Parents' experiences in accessing diagnostic and treatment services for children and adolescents diagnosed with autism spectrum disorder.

Francesco Saverio Sorrentino, Lorenzo Gardini, Luigi Fontana, Mutali Musa, Andrea Gabai, Antonino Maniaci, Salvatore Lavalle, Fabiana D'Esposito, Andrea Russo, Antonio Longo, Pier Luigi Surico, Caterina Gagliano, and Marco Zeppieri. 2024. Novel approaches for early detection of retinal diseases using artificial intelligence. *Journal of Personalized Medicine*, 14(7).

V. Pream Sudha and M. S. Vijaya. 2021. Recurrrent Neural Network Based Model for Autism Spectrum Disorder Prediction using Codon Encoding. *Journal of The Institution of Engineers (India): Series B*.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2023. Attention is all you need.