Autism Risk Classification using Graph Neural Networks Applied to Gene Interaction Data

1st Kyle Riccard
Department of Computer Science
Fairfield University
Fairfield, United States
Kyle.Riccardi@student.fairfield.edu

2nd Danushka Bandara
Department of Computer Science
Fairfield University
Fairfield, United States
dbandara@fairfield.edu

***Abstract*—We use a gene interaction network to predict which genes are associated with Autism Spectrum Disorder (ASD), thus allowing for earlier detection of ASD. ASD is a disorder that affects the development of the brain. No medical tests can diagnose ASD, so the diagnosis is based on the individual's behavior and development. We used gene location in the chromosome band and their possible interaction with each other to help classify the gene association with ASD. Graph sage and graph convolutional network (GCN) models were studied to organize these associations. The model's performance was analyzed using various parameters to obtain optimal performance. The graph sage model obtained an area under the ROC curve (AUC) of 0.85 percent on testing and 0.85 on validation. The GCN model achieved 0.78 AUC on testing and 0.78 AUC on validation. GCN achieves a specificity of 0.96 and a sensitivity of 0.63. Graph Sage achieved a specificity between 0.96 and a sensitivity of 0.94. Our classification performance demonstrates the capability of graph sage to classify genes carrying association risk for ASD. This analysis also demonstrates the inability of GCN to recognize association, as determined by the low sensitivity. Our results show that it is possible to apply graph neural networks to understand the link between diseases and genetics.**

***Index Terms*—Graph Convolutional Network (GCN), Graph Sage, Autism Spectrum Disorder (ASD)**

I. Introduction

With the widespread availability of biological data, many studies use genetic data to assess disease association [1]. For example, Feng et al. used biological data such as protein-protein interaction (PPI) networks and omic data to determine links between proteins that will help discover possible causations of diseases. Motsinger et al. [2] designed a genetic programming optimized neural network (GPNN), which identified gene interaction related to Parkinson's disease using gene-gene interaction networks. Li et al. [3] used a permutation-based random forest model to identify the top interaction of single nucleotide polymorphism pairs. Li et al. [4] used gene-by-gene networks to visualize the association with ASD. From the examples above, it is clear that biological data can allow us to understand the causes of diseases and helps us further understand root causes.

This study uses publicly available gene datasets to identify genes that cause autism. Autism is typically detected using behavioral factors [5]. However, genetic methods have recently been employed in autism detection. Some methods use genetic data from individuals to classify their risk for autism [6]. Another approach is to classify the genes that cause the risk for autism, which is the approach we take in this paper.

ASD risk has been shown to vary with genetics. Eighty percent of an individual's risk for developing autism is based on genetic makeup [7]–[9]. ASD risks are associated with copying number variants (CNV) or de novo mutations (DNM) [6], [10], [11]. CNV refers to a genome that sustains the performance of adding, removing, or copying segments of Deoxyribonucleic acid (DNA) [12]. The procedure required to create CNV may result in mutation, which leads to associated risk for ASD. DNM refers to genetic changes based on mutation sustained during both maturation of female gametes and division of stem cells within the trio (mother, father, and child) [13]. DNM is found using the genetics of the trio to determine possible mutations, which have a high inheritability [14]–[16]. ASD affects nearly 1-1.5 percent of people with a high inheritance rate from parent to child [17]–[20]. Furthermore, ASD affects 1 in 44 children who are eight years old (according to statistics from the USA in 2018) [21]. For example, 13.8 percent of a study considering 2270 ASD trios contained a de novo mutation [22]. The location of genes on a chromosome band (loci) is significant because there is a correlation between the location of genes and ASD [23], [24].

Many studies have attempted to classify topics associated with ASD. For instance, Colella et al. [25] used QuantiSNP, a hidden Markov model, to classify the mapping of CNV based on mapping genotypes. Furthermore, Sundaram et al. [26] designed DeepBipolar, a deep learning model that used genomic data to classify whether somebody is Bipolar or not. According to Rahman et al. [27], research must include features that help with the early identification and intervention of ASD to improve the efficiency of diagnosing ASD. These studies show the capability of machine learning models to demonstrate new findings and to identify whether a gene has shown a risk associated with ASD or not.

Our study demonstrates the use of graph neural network techniques to classify ASD risk association based on the possible location a gene can exist and its interaction network. [28] attempted to do the same using a Bayes network yielding an accuracy of 78.31 percent. Wang et al. [29] used Deep Autism, a convolutional neural network (CNN) model, to

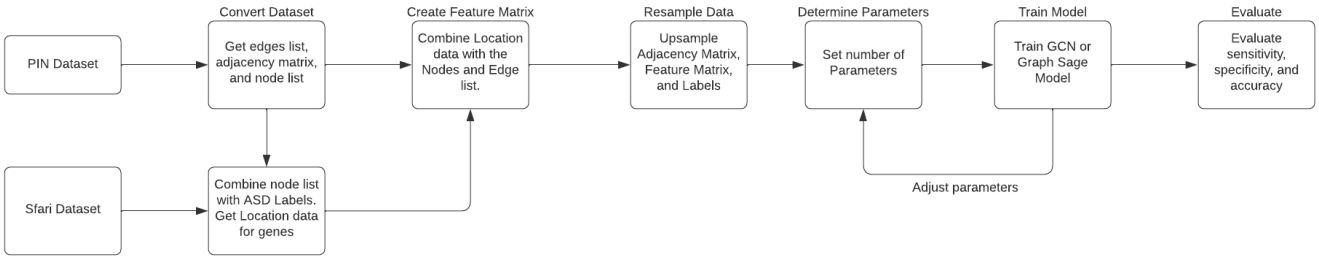

Fig. 1. Pipeline of the methodology

classify ASD risk, which achieved an accuracy of 88.60 percent. Beyreli et al. [30] used a multitask graph convolutional network system called DeepND to find an association between ASD gene scores and Intellectual disability (ID). DeepND uses risk associations and networks for specific disorders to identify the novel connections in ASD/ID. Our model uses the chromosome band feature and the graph network to classify risk associations in existing and novel genes outside the ASD gene list. The main contribution of our research is the application of Graph neural networks to the same problem. It results in a model that accurately classifies whether a gene contains risk association with ASD using graph neural network techniques.

II. Methodology

The gene network was generated using gene interaction data from two datasets. The analysis pipeline is shown in Figure 1. Since this is a previously untested dataset, different graph neural network models were tested to determine the performance variation. A multi-layer perceptron (MLP) model serves as a baseline for performance comparison.

A. Dataset

TABLE I
THE METADATA OF THE PROTEIN INTERACTION NETWORK (PIN) NETWORK WHILE ALSO SHOWING THE ASSOCIATION AND NON-ASSOCIATION ESTABLISHED BY THE SFARI DATASET.

Network Information	Network Parameters
nodes	12,216
edges	23,482
nodes not associated with ASD	11,403
nodes associated with ASD	813

We used two publicly available gene interaction datasets. (1) The PIN dataset is a mapping of genes and other genes that biochemically interact (Protein Binding, Promoter Binding, RNA Binding, Protein Modification, Direct Regulation, and Autoregulation) with them [31]. This dataset also contains 471 unique locations on chromosome bands. All of its interactions can allow us to see these genes' entire relationship with ASD. (2) The Sfari dataset [32] only contains genes that associate some evidence or strong evidence of association with ASD. The reports gathered on a gene increase the chance of a gene having a high-risk association with ASD. All genes within the Sfari dataset show a risk associated with ASD. The two datasets share 813 genes in common. These genes constitute our combined network. All Sfari dataset genes are given a label of one to establish their Risk association with ASD. In contrast, genes that don't exist in the Sfari dataset have been labeled a zero indicating that they do not correlate with ASD.

B. Data Extraction

Using the two datasets, we filtered out the gene interactions with no association between ASD in both genes. Also, any genetic interaction containing non-human genes was filtered out. The gene interaction was used to build an adjacency matrix shown in Figure 4 with genes as nodes and interactions as edges. Figure 3 shows the one-hot encoding used on the node features calculated from 471 chromosome band locations. These one-hot encodings represent how many locations on chromosome bands could contain this gene. Node labels represent the association/non-association of each node/gene with ASD.

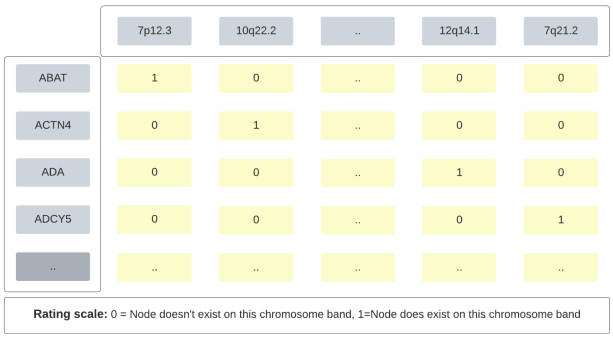

Fig. 3. Feature matrix of genes by their chromosome band location.

C. Preprocessing

The node labels were imbalanced, as seen in Table I between association/ non-association. Therefore, we used upsampling to balance the dataset. The upsampling algorithm randomly duplicates the minority class to even out the distribution between both classes. After resampling, the data was split 50 percent for training and the other 50 percent for validation and testing. This training, validation, and testing split was initially proposed in Zhao et al. study using GraphSmote, upsampling,

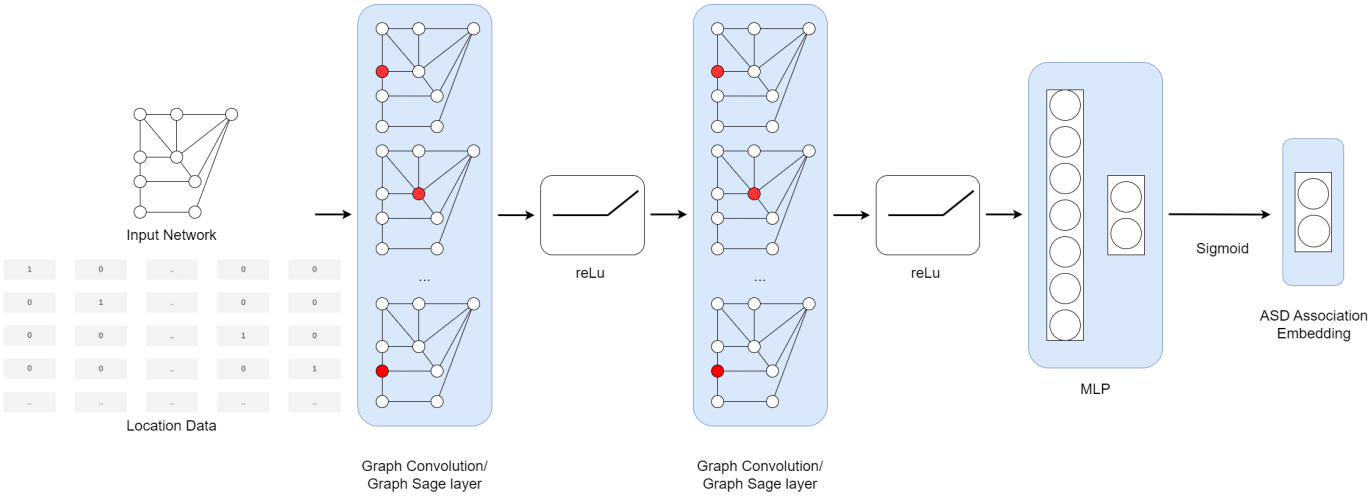

Fig. 2. Model Architecture of Graph Neural Network Model

and other re-sampling techniques for Graph neural networks [33].

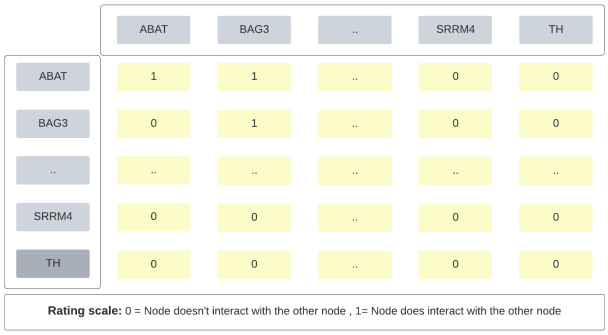

Fig. 4. Adjacency Matrix of the graph network.

D. Graph Convolutional Network

The model structure is shown in Figure 2. Graph convolutional network (GCN) is developed to use the convolutional properties on a graph to classify edges or nodes. Such a model can classify new connections or more information on nodes in genetic network datasets. The GCN model takes an input of a feature matrix, adjacency matrix, and weight matrix (the weight matrix is randomly generated). All three are multiplied together, returning a new feature matrix. This process called the propagation layer (1), is the essential part of a GCN model [34]. The GCN propgation layer takes adjacency matrix (A), feature matrix (H), and weight matrix (W).

$$f(H^l, A) = \sigma(AH^lW) \tag{1}$$

This layer allows us to use the graph structure and node features to accurately create node features containing information associated with the adjacency matrix. This process is done twice to encode the data and then decode using another propagation layer followed by a sigmoid function. The sigmoid function returns the embedding for classifying the association/non-association to ASD. There is a rectified linear unit (reLu) activation function between each layer.

E. Graph Sage

The Graph Sage model is similar to GCN using the Figure 2 model structure; however, this model takes an adjacency matrix, node feature, and labels to do a different process for determining a new feature matrix. Graph Sage takes a sample of neighboring connections within the network to aggregate those sampled features into a singular feature matrix. This process aggregates (2) the features of all neighboring features with the original node to create new features for that node [35].

$$h^k_{N(v)} \leftarrow AGGREGATE_k(h^{k-1}_u, \forall u \in N(v)) \tag{2}$$

This aggregation is done for every node until it is aggregated with its neighbors. Once the aggregated features are collected, they are concatenated with the original node features (3), creating a new feature matrix.

$$h^k_v \leftarrow \sigma(W^k * CONCAT(h^{k-1}_v, h^k_{N(v)})) \tag{3}$$

This process is done twice for encoding and classifying, which is then fed through a decoder, which uses a fully connected layer followed by a sigmoid function to create the embedding matrix used for classification.

F. Baseline

The baseline is determined by exempting the network and replacing the two graph neural network layers with MLP layers. This baseline allows us to determine the network's effect on all our evaluating parameters.

III. Experiment

TABLE II
PARAMETERS USED FOR TRAINING THE MODELS.

Parameter Name	Value
Learning rate	0.0005
Weight decay	5e-4
Batch size	64
Epoch	5000
Output Function	Sigmoid
Activation Function	reLu
Optimizer	Adam

The models were created using Python [36] and the PyTorch library [37]. On each cross-validation split, we train for 5000 epochs (without early stopping) using the Adam optimizer with a learning rate of 0.0005. Other hyper-parameters are chosen as follows: 5e-4 (L2 regularization, first layer), n (number of units for each hidden layer), and 0.0005 (learning rate). Results are summarized in Table II [38].

IV. Results

The evaluation parameters used for this experiment are units per layer, accuracy, specificity, sensitivity, and area under the curve (AUC) for the test dataset. For the validation dataset, we evaluate it based on AUC.

A. Validation AUC Evaluation

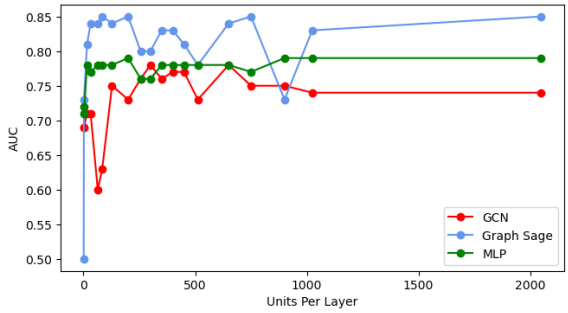

Fig. 5. Variation of validation AUC by the units per layer.

The models show in Figure 5 that AUC for GCN peaks at 0.78 at 300 units per layer. The AUC value tapers off around 400 units per layer with an AUC of 0.75. The peak for Graph Sage is 0.85 at 84 units per layer. The Graph Sage's tapering off can be seen at 400 units per layer with an AUC of 0.83. Our AUC for MLP at 400 units per layer is 0.78.

B. Test Accuracy Evaluation

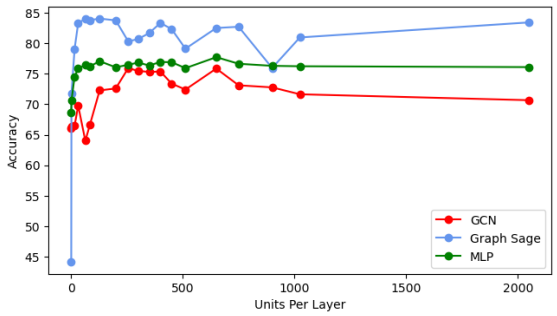

Fig. 6. Variation of test accuracy(%) by the units per layer.

Figure 6 shows GCN peak accuracy at 128 units per layer and Graph Sage at 650 units per layer. Graph Sage has a peak accuracy of 84.03 percent and a GCN peak accuracy of 75.91. The MLP peak accuracy of 77.71 is at 650 units per layer. The test accuracy drops off in improvement after 128 units per layer.

C. Test Specificity Evaluation

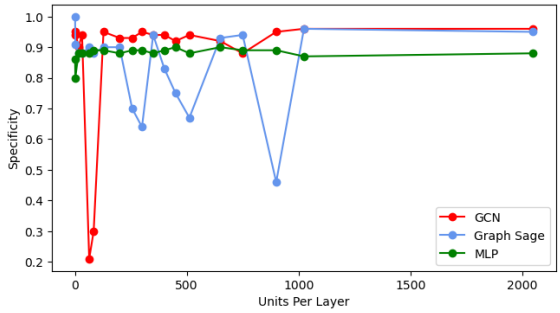

Fig. 7. Variation of test specificity by the units per layer.

The performance of these models shows in Figure 7 a specificity peak at 128 units per layer for GCN and 450 units per layer for Graph Sage. Graph Sage had a maximum specificity of 0.96 and GCN specificity of 0.96. Our baseline peaks at 650 units per layer with a specificity of 0.90.

D. Test Sensitivity Evaluation

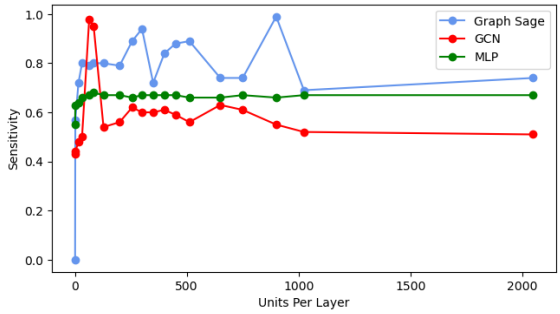

Fig. 8. Variation of test sensitivity by the units per layer.

TABLE III
THE PEAK PERFORMANCE ON OUR EVALUATION PARAMETERS USED ON THE VALIDATION AREA UNDER THE CURVE (AUC), TESTING DATASET (ACCURACY(%), TESTING AREA UNDER THE CURVE (AUC), TESTING SENSITIVITY, AND TESTING SPECIFICITY).

Model Type	Validation AUC	Test AUC	Test Specificity	Test Sensitivity	Test Accuracy (%)
Graph Sage	0.85	0.85	0.96	0.94	84.03
GCN	0.78	0.78	0.96	0.63	75.91
MLP	0.79	0.78	0.90	0.67	77.71

The models show in Figure 8 a sensitivity peak for Graph Sage at 1024 units per layer and GCN at 650 units per layer. Graph Sage had a maximum sensitivity of 0.96, and GCN had a Sensitivity of 0.63. GCN and Graph Sage start with high sensitivity for our analysis due to extremely low specificity. Those peaks would suffice for evaluating the model due to failure in other areas. Our baseline has a peak sensitivity of 0.77 at 84 units per layer.

E. Test AUC Evaluation

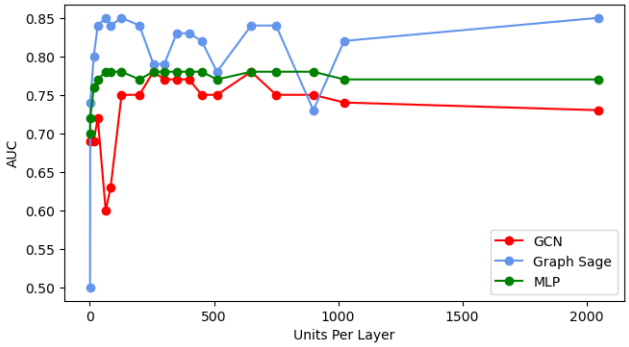

Fig. 9. Variation of test sensitivity by the units per layer.

The models show in Figure 9 an AUC peak of 0.85 at 2048 units per layer for Graph Sage and a GCN peak of 0.78 at 650. Our baseline AUC peak is 0.78 at 64 units per layer. The AUC for these models drops off with increased improvement at 300 units per layer.

V. DISCUSSION

This experiment demonstrates a graph representation of gene interactions for autism risk detection. We also show that graph neural networks are suitable for classifying this network data. According to Table III, the graph sage model had the best accuracy of 84.03 percent on the test set. Furthermore, it achieved an AUC of 0.85 for testing and 0.85 for validation at 2048 units per layer. It had a false negative rate of 26 percent and a false positive rate of 5 percent at 2048 units per layer. Showing a false positive rate close to the peak, which would be 4, but a false negative rate, which is nearly 20 percent off the peak of 0.94. GCN interestingly achieved an evaluation lower than MLP at Validation and Test AUC of 0.78. A false negative rate of 38 percent and a false positive rate of 4 percent. The model shows its ability to classify true positives while demonstrating a higher tendency to misclassify true negatives. Compared to the baseline MLP, Graph Sage performs better in all case. On the other hand, GCN shows better performance on false positives, while MLP performs better on false negatives. The MLP network on validation AUC only got 0.79; on testing, AUC got 0.78. Furthermore, specificity obtained 0.90, and sensitivity received 0.67. Lastly, the test accuracy was 77.71.

The low false positive and false negative rates indicate the viability of the graph sage method for diagnosis purposes. This would imply that if a patient has one or more genes associated with autism risk, they are at greater risk for the disease. The advantage of this method of diagnosis is that we can detect autism at very early stages and provide proper treatment and care to the individual.

GCN shows a performance comparable to MLP, showing that this approach isn't sufficient for an attempt to analyze this problem. The inductive representation of graph sage shows a better approach for attempting to understand the genetic relationship, which is missing from the GCN approach for this project.

The limitation of our method is that it only does a binary classification between autism risk/ no risk. It does not consider the different risk levels available in the Sfari dataset. Further studies should consider this factor. Also, our method doesn't consider any epigenetic factors. However, our method can be an indicator that can be combined with other indicators to come up with a holistic method for an autism diagnosis.

VI. CONCLUSION

We demonstrate the use of graph neural network models to identify ASD association based on ASD Loci and a gene interaction network to diagnose ASD earlier and find possible causation by linking more genes to ASD. We achieved 84 percent accuracy using our model, which shows promise that classifying the association risk of these genes using graph neural networks is possible. We also demonstrate gene location in the chromosome bands as our node features. Our method can be part of a diagnostic process for autism. Furthermore, understanding the possible causation for these mutations will help determine the root causes of autism. Our results show that graph neural networks are promising for understanding the relationship between genes and diseases.

REFERENCES

[1] J. Feng, A. Zeng, Y. Chen, P. Payne, and F. Li, "Signaling interaction link prediction using deep graph neural networks integrating protein-protein interactions and omics data," *bioRxiv*, 2020. [Online]. Available: https://www.biorxiv.org/content/early/2020/12/24/2020.12.23.424230

[2] L. S. M. G. e. a. Motsinger, A.A., "Gpnn: Power studies and applications of a neural network method for detecting gene-gene interactions in studies of human disease," *BMC Bioinformatics*, vol. 7, no. 39, 2006. [Online]. Available: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-7-39

[3] M. J. A. A. e. a. Li, J., "Detecting gene-gene interactions using a permutation-based random forest method," *BioData Mining*, vol. 9, no. 14, 2016. [Online]. Available: https://biodatamining.biomedcentral.com/articles/10.1186/s13040-016-0093-5

[4] J. Li, M. Shi, Z. Ma, S. Zhao, G. Euskirchen, J. Ziskin, A. Urban, J. Hallmayer, and M. Snyder, "Integrated systems analysis reveals a molecular network underlying autism spectrum disorders," *Molecular Systems Biology*, vol. 10, no. 12, p. 774, 2014. [Online]. Available: https://www.embopress.org/doi/abs/10.15252/msb.20145487

[5] F. Happe, A. Ronald, and R. Plomin, "Time to give up on a single explanation for autism," *Nature Neuroscience*, vol. 9, no. 10, pp. 1218 – 1220, Oct. 2006.

[6] Z. D. C. G. E. I. P. C. Skafidas E, Testa R, "Predicting the diagnosis of autism spectrum disorder using gene pathway analysis," *Mol Psychiatry*, vol. 19, no. 504-510, 2014.

[7] Autism Spectrum Disorder Working Group of the Psychiatric Genomics Consortium, BUPGEN, Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, 23andMe Research Team, J. Grove, S. Ripke, T. Als, M. Mattheisen, R. Walters, H. Won, J. Pallesen, E. Agerbo, O. Andreassen, R. Anney, S. Awashti, R. Belliveau, F. Bettella, J. Buxbaum, J. Bybjerg-Grauholm, M. Bækvad-Hansen, F. Cerrato, K. Chambert, J. Christensen, C. Churchhouse, K. Dellenvall, D. Demontis, S. De Rubeis, B. Devlin, S. Djurovic, A. Dumont, J. Goldstein, C. Hansen, M. Hauberg, M. Hollegaard, S. Hope, D. Howrigan, H. Huang, C. Hultman, L. Klei, J. Maller, J. Martin, A. Martin, J. Moran, M. Nyegaard, T. Nærland, D. Palmer, A. Palotie, C. Pedersen, M. Pedersen, F. Goes, D. MacKinnon, F. Mondimore, J. DePaulo, and J. Potash, "Identification of common genetic risk variants for autism spectrum disorder," *Nature Genetics*, vol. 51, no. 3, pp. 431–444, Mar. 2019, publisher Copyright: © 2019, The Author(s), under exclusive licence to Springer Nature America, Inc.

[8] H. Wang and P. Avillach, "Diagnostic classification and prognostic prediction using common genetic variants in autism spectrum disorder: Genotype-based deep learning," *JMIR Med Inform*, vol. 9, no. 4, p. e24754, Apr 2021. [Online]. Available: https://medinform.jmir.org/2021/4/e24754

[9] H. Wang, L. Li, L. Chi, and Z. Zhao, "Autism screening using deep embedding representation," in *Computational Science – ICCS 2019*, J. M. F. Rodrigues, P. J. S. Cardoso, J. Monteiro, R. Lam, V. V. Krzhizhanovskaya, M. H. Lees, J. J. Dongarra, and P. M. Sloot, Eds. Cham: Springer International Publishing, 2019, pp. 160–173.

[10] J. Vorstman, J. Parr, D. Moreno-De-Luca, R. Anney, J. Nurnberger, and J. Hallmayer, "Autism genetics: opportunities and challenges for clinical translation," *Nature Reviews Genetics*, vol. 18, 03 2017.

[11] T. L. Gaugler, L. Klei, S. J. Sanders, C. Bodea, A. P. Goldberg, A. B. Lee, M. C. Mahajan, D. Manaa, Y. Pawitan, J. G. Reichert, S. Ripke, S. Sandin, P. Sklar, O. Svantesson, A. Reichenberg, C. M. Hultman, B. Devlin, K. Roeder, and J. D. Buxbaum, "Most genetic risk for autism resides with common variation," *Nature genetics*, vol. 46, pp. 881 – 885, 2014.

[12] T. Shaikh, "Copy number variation disorders," *Current Genetic Medicine Reports*, vol. 5, no. 183-190, 2017.

[13] K. Samocha, E. Robinson, S. Sanders, C. Stevens, A. Sabo, L. McGrath, J. Kosmicki, K. Rehnström, S. Mallick, A. Kirby, D. Wall, D. MacArthur, S. Gabriel, M. DePristo, S. Purcell, A. Palotie, E. Boerwinkle, J. Buxbaum, E. Cook, and M. Daly, "A framework for the interpretation of de novo mutation in human disease," *Nature genetics*, vol. 46, 08 2014.

[14] I. Iossifov, B. O'Roak, S. Sanders, M. Ronemus, N. Krumm, D. Levy, H. Stessman, K. Witherspoon, L. Vives, K. Patterson, J. Smith, B. Paeper, D. Nickerson, J. Dea, S. Dong, L. Gonzalez, J. Mandell, S. Mane, M. Murtha, and M. Wigler, "The contribution of de novo coding mutations to autism spectrum disorder," *Nature*, vol. 515, 10 2014.

[15] J. Veltman and H. Brunner, "De novo mutations in human genetic disease," *Nature reviews. Genetics*, vol. 13, pp. 565–75, 07 2012.

[16] J. Gratten, P. Visscher, B. Mowry, and N. Wray, "Interpreting the role of de novo protein-coding mutations in neuropsychiatric disease," *Nature genetics*, vol. 45, pp. 234–8, 03 2013.

[17] L. Rylaarsdam and A. Guemez-Gamboa, "Genetic causes and modifiers of autism spectrum disorder," *Frontiers in Cellular Neuroscience*, vol. 13, p. 385, 08 2019.

[18] A. Thapar and M. Rutter, "Genetic advances in autism," *Journal of Autism and Developmental Disorders*, vol. 51, pp. 1–12, 12 2021.

[19] H. Wang, L. Chi, H. Yang, L. Li, and Z. Zhao, *A deep learning predictive classifier for autism screening and diagnosis*, 01 2021, pp. 215–236.

[20] S. Sandin, P. Lichtenstein, R. Kuja-Halkola, C. M. Hultman, H. Larsson, and A. Reichenberg, "The heritability of autism spectrum disorder," *JAMA*, vol. 318, p. 1182, 09 2017.

[21] M. Maenner, K. Shaw, A. Bakian, D. Bilder, M. Durkin, A. Esler, S. Furnier, L. Hallas-Muchow, J. Hall-Lande, A. Hudson, M. Hughes, M. Patrick, K. Pierce, J. Poynter, A. Salinas, J. Shenouda, A. Vehorn, Z. Warren, J. Constantino, and M. Cogswell, "Prevalence and characteristics of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, united states, 2018," *Morbidity and mortality weekly report. Surveillance summaries (Washington, D.C. : 2002)*, vol. 70, pp. 1–16, 12 2021.

[22] S. De Rubeis, X. He, A. Goldberg, C. Poultney, K. Samocha, A. E. Cicek, Y. Kou, l. Liu, M. Fromer, S. Walker, T. Singh, L. Klei, J. Kosmicki, S. Fu, B. Aleksic, M. Biscaldi-Schäfer, P. Bolton, J. Brownfeld, J. Cai, and J. Buxbaum, "Synaptic, transcriptional and chromatin genes disrupted in autism," *Nature*, 10 2014.

[23] M. Ronemus, I. Iossifov, D. Levy, and M. Wigler, "The role of de novo mutations in the genetics of autism spectrum disorders," *Nature reviews. Genetics*, vol. 15, pp. 133–41, 02 2014.

[24] S. Sanders, X. He, A. Willsey, A. Ercan-Sencicek, K. Samocha, A. E. Cicek, M. Murtha, V. Bal, S. Bishop, S. Dong, A. Goldberg, C. Jinlu, J. Keaney III, L. Klei, J. Mandell, D. Moreno De Luca, C. Poultney, E. Robinson, L. Smith, and M. State, "Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci," *Neuron*, vol. 87, pp. 1215–1233, 09 2015.

[25] S. Colella, C. Yau, J. Taylor, G. Mirza, H. Butler, P. Clouston, A. Bassett, A. Seller, C. Holmes, and J. Ragoussis, "Quantisnp: an objective bayes hidden-markov model to detect and accurately map copy number variation using snp genotyping data," *Nucleic acids research*, vol. 35, pp. 2013–25, 02 2007.

[26] S. Laksshman, R. Bhat, V. Viswanath, and X. Li, "Deepbipolar: Identifying genomic mutations for bipolar disorder via deep learning," *Human mutation*, vol. 38, 06 2017.

[27] M. M. Rahman, U. Lateef, R. Muniyandi, S. Sahran, S. Mohamed, and R. a razak, "A review of machine learning methods of feature selection and classification for autism spectrum disorder," *Brain Sciences*, vol. 10, p. 949, 12 2020.

[28] M. Gök, "A novel machine learning model to predict autism spectrum disorders risk gene," *Neural Computing and Applications*, vol. 31, 10 2019.

[29] H. Wang and P. Avillach, "Genotype-based deep learning in autism spectrum disorder: Diagnostic classification and prognostic prediction using common genetic variants (preprint)," 10 2020.

[30] Beyreli, O. Karakahya, and A. E. Cicek, "Deepnd: Deep multitask learning of gene risk for comorbid neurodevelopmental disorders," *Patterns*, vol. 3, p. 100524, 06 2022.

[31] W. Pereanu, E. Larsen, I. Das, M. Estevez, A. Sarkar, S. Spring-Pearson, R. Kollu, S. Basu, and S. Banerjee-Basu, "Autdb: A platform to decode the genetic architecture of autism," *Nucleic acids research*, vol. 46, 11 2017.

[32] B. Abrahams, D. Arking, D. Campbell, H. Mefford, E. Morrow, L. Weiss, I. Menashe, T. Wadkins, S. Banerjee-Basu, and A. Packer, "Sfari gene 2.0: A community-driven knowledgebase for the autism spectrum disorders (asds)," *Molecular autism*, vol. 4, p. 36, 10 2013.

[33] T. Zhao, X. Zhang, and S. Wang, "Graphsmote: Imbalanced node classification on graphs with graph neural networks," 03 2021.

[34] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *International Conference on Learning Representations*, 2017. [Online]. Available: https://openreview.net/forum?id=SJU4ayYgl

[35] W. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," 06 2017.

[36] G. Van Rossum and F. L. Drake Jr, *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

[37] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner,

L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019, pp. 8024–8035. [Online]. Available: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

[38] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations*, 12 2014.