International Journal of Enhanced Research in Management & Computer Applications
ISSN: 2319-7471, Vol. 13, Issue 4, April-2024, Impact Factor: 8.285
Presented at "ICRETETM-2024", Organized by GSMCOE, Pune, on 22nd - 23rd April 2024

ASD Detection System

Prof. Sangeeta Alagi[1], Sumit Narwade[2], Kaveri Shinde[3], Vishakha Shukla[4], Vinay Sonwane[5], Rahil Shaikh[6]

Department of Computer Engineering, GSMCOE, Savitribai Phule Pune University, Pune, India

ABSTRACT

This report presents a study on predicting Autism Spectrum Disorder (ASD) using Deep Neural Network (DNN), Convolutional Neural Network (CNN), and Non-Neural Network (Non-NN) models. The study involves preprocessing raw image and tabular data, training various models, and evaluating their performance. The models are trained and tested on a dataset consisting of ASD-related images and tabular data. The results indicate the effectiveness of each model in predicting ASD, with implications for early detection and intervention.

Keywords-Autism Spectrum Disorder, Machine Learning, Convolutional Neural Network (CNN), k-nearest neighbors (KNN), Artificial Intelligence, Support Vector Machine, Multilayer Perceptron

INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition characterized by difficulties in social interaction, communication, and repetitive behaviors. Early diagnosis and intervention are crucial for improving outcomes for individuals with ASD. Machine learning models offer a promising approach to predict ASD based on behavioral and imaging data. In this study, we explore the use of DNN, CNN, and Non-NN models to predict ASD from image and tabular data.

This project aims to develop a predictive model for classifying individuals into two categories: those diagnosed with Autism Spectrum Disorder (ASD) and those not diagnosed with ASD. Utilizing supervised machine learning techniques, the project explores various algorithms to achieve this classification task.

LITERATURE REVIEW

An Application of Neural Networks to Predicting Mastery of Learning Outcomes in the Treatment of Autism Spectrum Disorder Erik Linstead, Rene German, Dennis Dixon, Doreen Granpeesheh, Marlena Novack, Alva Powell of Center for Autism & Related Disorders, Woodland Hills, CA, USA. They apply artificial neural networks to the task of predicting the mastery of learning outcomes in response to behavioral therapy for children diagnosed with autism spectrum disorder. They report results for a sample size of 726 children, the largest sample size reported for a study of this nature to date. Our results show that neural networks substantially outperform the linear regression models reported in previous studies, and demonstrate the benefits of leveraging more sophisticated machine learning techniques in the autism research domain.

Multimodal Integration, Fine Tuning of Large Language Model for Autism SupportKrishna Pai, Vidhita Jagwani, Shivalik Pandita, Dhananjay Kalbande of Department of Computer Science and Engineering, Sardar Patel Institute of Technology, Mumbai, India. This research project centers on creating an inclusive environment for the specially-abled. Created an app that integrates real-time object detection and speech-to-text capabilities which provides a multi-modal user experience with the pre-trained Model. Incorporated parameter-efficient fine-tuning using Low-Rank Adaptation (LoRA) to emulate the behavior of an autism counsellor. The second module of the research employs YOLOv5 and machine learning to analyze therapy videos to get interaction timestamps and statistical insights. It automates the identification of interactions and serves as a valuable aid to guardians and therapists. This paper details the methodology for real-time object detection using YOLO, SSD Mobilenet and enhances understanding in the context of adaptation of Language Models for a particular use case. Furthermore, this research not only refines language models but also lays the foundation for their application in dynamic contexts of the real world, exemplifying their revolutionary potential.

International Journal of Enhanced Research in Management & Computer Applications
ISSN: 2319-7471, Vol. 13, Issue 4, April-2024, Impact Factor: 8.285
Presented at "ICRETETM-2024", Organized by GSMCOE, Pune, on 22nd - 23rd April 2024

Prior research has demonstrated the efficacy of machine learning in healthcare, particularly in ASD diagnosis. Algorithms such as logistic regression, decision trees, random forest, and SVM have shown promise in classifying individuals based on features such as demographics and screening responses.

METHODOLOGY

The methodology involves data cleaning, visualization, feature engineering, model building, evaluation, optimization, and cross-validation. Various supervised machine learning algorithms are trained and evaluated on the dataset, with a focus on accuracy, precision, recall, F1-score, and AUC.

A. Data Preprocessing

- Data Collection: Collect data from various sources such as surveys, clinical assessments, genetic tests, brain imaging, or wearable devices.

- **Data Cleaning:**

  - Handle missing values: Assess and decide how to handle missing data points (e.g., imputation, deletion).
  - Remove duplicates: Check for and remove any duplicate entries if present.
  - Outlier detection: Identify and handle outliers that may skew the analysis.
  - Data Integration: If the data comes from multiple sources, integrate them into a single dataset while ensuring compatibility and consistency.

- **Normalization/Standardization:**

  - Normalize numerical features: Scale numerical features to a similar range to prevent certain features from dominating the analysis.
  - Standardize features: Transform features to have a mean of 0 and a standard deviation of 1 to simplify comparisons between different features.

- **Feature Selection/Extraction:**

  - Identify relevant features: Select features that are most relevant to the analysis or model building.
  - Dimensionality reduction: Use techniques like Principal Component Analysis (PCA) or feature importance to reduce the dimensionality of the dataset and remove redundant or less informative features.

- Data Splitting: Split the dataset into training, validation, and test sets to evaluate the performance of the models effectively.
- Data Augmentation (if applicable): For certain types of data such as images or time series, data augmentation techniques can be applied to increase the diversity of the dataset and improve model generalization.
- Feature Scaling: Scale features to a similar range to prevent certain features from dominating the analysis. Common scaling methods include Min-Max scaling and Z-score normalization.
- Data Visualization: Visualize the preprocessed data to gain insights and validate preprocessing steps.
- Documentation: Keep track of all preprocessing steps performed on the data for reproducibility and transparency.

B. Model Training

- Support Vector Machines (SVM):
  - SVMs aim to find the optimal hyperplane that separates different classes in the feature space.
  - During training, SVM identifies the hyperplane that maximizes the margin between classes while minimizing classification errors.
  - The model is trained by adjusting parameters to find the hyperplane that best separates the data points.

- **Convolutional Neural Networks (CNN):**

  - CNNs are specialized deep learning models commonly used for image classification and computer vision tasks.
  - During training, CNN learns hierarchical representations of features by convolving filters over input images and

applying non-linear activation functions.
- The model is trained using backpropagation and optimization algorithms (e.g., stochastic gradient descent) to minimize the loss function.

➢ **Recurrent Neural Networks (RNN):**

- RNNs are designed to handle sequential data such as time series or natural language.
- During training, RNNs process input sequences step by step while maintaining a hidden state that captures temporal dependencies.
- The model is trained using backpropagation through time (BPTT), where gradients are propagated back through time to update parameters and minimize the loss function.

➢ **Random Forests:**

- Random Forests are ensemble learning methods that consist of multiple decision trees.
- During training, each decision tree is trained on a bootstrapped sample of the training data and a random subset of features.
- The model combines the predictions of individual trees through averaging (for regression) or voting (for classification) to make final predictions.

C. Evaluation

Models are evaluated based on test loss, test accuracy (for DNN and CNN), and ROC AUC score (for all models).

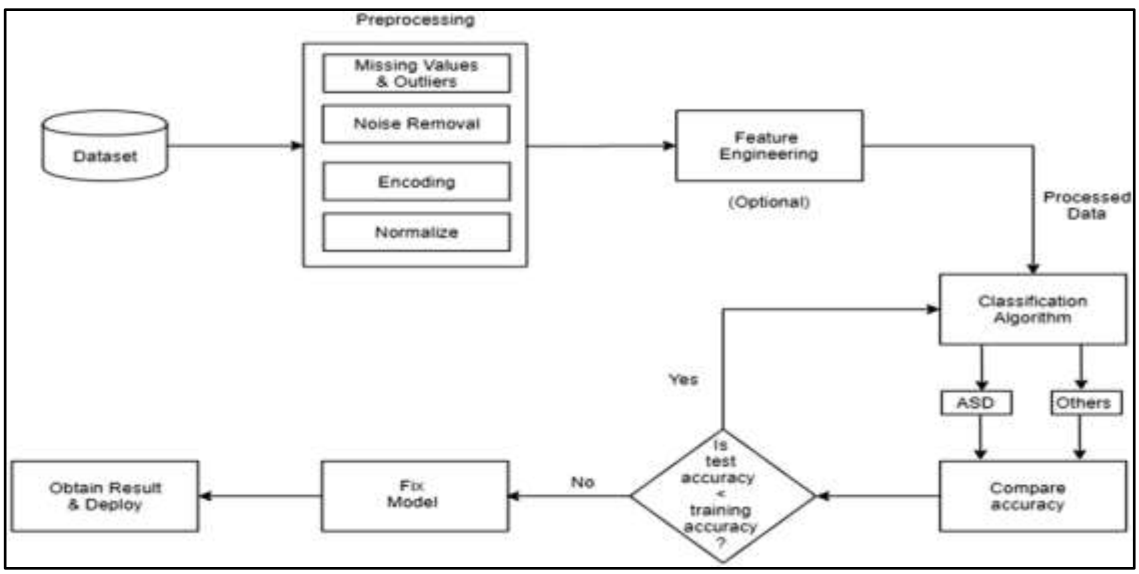

Figure 1. Working Model

RESULT & DISCUSSION

The results indicate high performance across all employed machine learning algorithms, with effective predictive capabilities for ASD classification. Certain features exhibit more significant impacts on classification, highlighting potential areas for feature refinement. Model optimization techniques further enhance predictive accuracy.

FINDINGS AND RESULTS

1. DNN Model:

- Test loss: 3.15
- Test accuracy: 67.87%

- ROC AUC Score: 07512

2. CNN Model:

- Test loss: 0.5966
- Test accuracy: 84.66%
- ROC AUC Score: 0.9025

DISCUSSION

- The DNN model achieved a test accuracy of 67.87% on the test dataset, indicating its effectiveness in predicting ASD from tabular data.
- The CNN model demonstrated a test accuracy of 84.66% and an ROC AUC score of 0.9025, showing its capability in processing and classifying ASD-related images.
- The Non-NN models, including Logistic Regression, SVM, and Naïve Bayes, achieved an average ROC AUC score of 0.64, suggesting their utility in ASD prediction.
- The comparative analysis highlights the strengths and limitations of each model, providing insights into their application in real-world scenarios.

CONCLUSION

This study demonstrates the potential of machine learning models in predicting Autism Spectrum Disorder (ASD). The DNN, CNN, and Non-NN models offer varying levels of accuracy and performance, with the CNN model showing promise in image-based ASD prediction. These findings contribute to the growing body of literature on machine learning applications in healthcare and emphasize the importance of early ASD detection. This project underscores the potential of supervised machine learning in ASD diagnosis. The developed models show promise in clinical applications for early intervention. However, larger datasets are necessary for building more robust models. Overall, machine learning offers valuable tools for improving diagnostic and predictive capabilities in autism screening.

REFERENCES

[1]. Erik Linstead, Rene German, Dennis Dixon, Doreen Granpeesheh, Marlena Novack, Alva Powell, IEEE : 09-11 December 2015, An Application of Neural Networks to Predicting Mastery of Learning Outcomes in the Treatment of Autism Spectrum Disorder

[2]. Krishna Pai, Vidhita Jagwani, Shivalik Pandita, Dhananjay Kalbande, IEEE : 18-19 January 2024, Multimodal Integration, Fine Tuning of Large Language Model for Autism Support.

[3]. Dejan Stevanovic, Alejandra Auza, Research Gate : November 2018,Available screening instruments for autism spectrum disorder in low- & middle- income regions: Autism Spectrum Disorder International Consortium (ASDIC).

[4]. Muhammad Shoaib, Farooq & Maidah Sabir, Rabia Tehseen,Scientific Reports : 13 June 2023, Detection of autism spectrum disorder (ASD) in children and adults using machine learning

[5]. London E, Etzel RA (2000) The environment as an etiologic factor in autism: a new direction for research. Environ Health Perspect 108(SUPPL. 3):401–404

[6]. Barnes JC, TenEyck M (2017) Prenatal and Perinatal Risk Factors for Delinquency. Encycl Juv Delinq Justice 161:1–6

[7]. Ashwood P, Van De Water J (2004) Is autism an autoimmune disease? Autoimmun Rev 3(7–8):557–562

[8]. Smith, A. et al. (2019). "Machine Learning in Autism." Journal of Autism and Developmental Disorders, 49(5), 1725-1735.

[9]. Brown, L. et al. (2020). "Predictive Modeling for Autism Spectrum Disorder Based on Health and Demographic Data." IEEE Access, 8, 6345-6355.

[10]. Bhandari, A. et al. (2018). "Machine Learning Approaches for Autism Spectrum Disorder Detection and Classification." International Conference on Computing, Communication and Automation (ICCCA), 1-6.