**Autism Detection using Machine Learning**

-Manohar B

**Autism Detection Using Machine Learning**

**Project Report**

**Bachelor of Technology**

**(Computer Science Engineering)**

**Submitted to**

****

**LOVELY PREOFESSIONAL UNIVERSITY**

**PHAGWARA, PUNJAB**

**SUBMITTED BY**

**Name of the Student: Manohar Bethu**

**Registration Number: 12100789**

**Annexure-II**

**To whom so ever it may concern**

I, **Manohar Bethu, 12100789**, hereby declare that the work done by me on “**Autism Detection using Machine Learning**”, is a record of original work for the partial fulfillment of the requirements for the award of the degree, Bachelor of Technology.

**Name of the Student:** Manohar Bethu

**Registration Number:** 12100789

**ACKNOWLEDGEMENT**

I would like to express my special gratitude to the teacher and instructor of the course Machine Learning who provided me with the golden opportunity to learn a new technology.

I would like to thank my college Lovely Professional University for offering such a course which not only improves my programming skills, but also taught me other new technology.

Then I would like to thank my parents and friends who have helped me with their valuable suggestions and guidance for choosing this course.

Finally, I would like to thank everyone who has helped me a lot.

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **S. No.** | **Contents** | **Page** |
| 1 | TITLE | 01 |
| 2 | STUDENT DECLARATION | 02 |
| 3 | ACKNOWLEDGEMENT | 03 |
| 4 | ABSTRACT | 05 |
| 5 | OBJECTIVE | 05 |
| 6 | INTRODUCTION | 06 |
| 7 | THEORETICAL BACKGROUND | 06-07 |
| 8 | LIBRARIES | 08 |
| 9 | FEATURE SET EXPLORATION | 09 |
| 10 | PREPARING THE DATA | 09 |
| 11 | DROPPING MISSING VALUES | 10 |
| 12 | VISUALIZATION | 10-12 |
| 13 | DATA PREPROCESSING | 12 |
| 14 | TRAINING AND TEST DATA | 14 |
| 15 | RANDOM FOREST CLASSIFICATION | 14 |
| 16 | DECISION MAKING TREE | 14 |
| 17 | SUPPORT VECTOR MACHINE | 15 |
| 18 | K-NEAREST NEIGHBORS | 15 |
| 19 | NAIVE BAYES | 16 |
| 20 | LOGISTIC REGRESSION | 16 |
| 21 | MODEL TUNING | 16-17 |
| 22 | CONCLUSION | 17 |

**ABSTRACT**

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder that often goes undiagnosed due to the lack of early and accurate detection methods. This research aims to develop a machine learning model that can effectively detect autism in children using a combination of behavioral, cognitive, and demographic data. By leveraging advanced machine learning techniques, we aim to improve the early identification of ASD, enabling timely intervention and support for affected individuals. This research will contribute to the development of more efficient and accessible diagnostic tools for ASD.

**OBJECTIVE**

Autism Spectrum Disorder (ASD) is a complex neurological condition that affects an individual’s social interactions, communication skills, and behavior. Traditional diagnostic methods, which rely on clinical assessments and standardized tests, can be time-consuming and costly. To address these challenges, machine learning techniques are being increasingly utilized to enhance the accuracy and efficiency of ASD detection. By analyzing various data inputs such as behavioral patterns, facial features, and genetic information, machine learning models can identify subtle indicators of autism that might be missed by conventional methods12.

Several machines learning algorithms, including Support Vector Machines (SVM), Random Forest Classifiers (RFC), and Convolutional Neural Networks (CNN), have shown promise in detecting ASD. These models can process large datasets to uncover patterns and correlations that are indicative of autism. For instance, deep learning models have been used to analyze facial images to detect autism with high accuracy2. The integration of machine learning in autism detection not only accelerates the diagnostic process but also reduces the burden on healthcare systems, making early intervention more accessible to those in need

**INTRODUCTION**

Introduction Autism Spectrum Disorder (ASD) is a developmental condition characterized by challenges in social interaction, communication, and repetitive behaviors. Early and accurate detection of ASD is crucial for providing timely interventions that can significantly improve the quality of life for individuals affected by this condition. Traditional diagnostic methods, which often involve lengthy clinical assessments and observations, can be both time-consuming and costly. To address these challenges, researchers are increasingly turning to machine learning techniques to enhance the detection and diagnosis of ASD.

Machine learning algorithms can analyze vast amounts of data to identify patterns and indicators of autism that might be overlooked by human evaluators. These algorithms utilize various data sources, including behavioral assessments, genetic information, and even facial recognition technology, to detect subtle signs of ASD. By leveraging the power of machine learning, it is possible to develop more efficient, accurate, and accessible diagnostic tools, ultimately leading to earlier diagnosis and better outcomes for individuals with autism

**THEORETICAL BACKGROUND**

1. What is **Machine Learning?**

Machine learning is a branch of artificial intelligence (AI) that focuses on the development of algorithms and statistical models that allow computers to learn from and make predictions or decisions based on data. Instead of explicitly programming a computer to perform a specific task, machine learning enables it to learn patterns and relationships within data to improve its performance over time.

1. Random Forest Classification

It constructs multiple decision trees during training and outputs the class that is the mode (classification) or mean (regression) of the predictions from individual trees It combines the predictions from multiple trees to improve generalization and robustness over a single decision tree. Random Forests help mitigate overfitting, improve prediction accuracy, and handle both classification and regression problems.

1. Decision Tree Classifier

A Decision Tree Classifier is a versatile machine learning algorithm that excels at making predictions based on a tree-like structure. This structure comprises nodes, where each internal node represents a decision point based on a specific feature. These decisions branch out, leading to further nodes or leaf nodes. Leaf nodes, the terminal points of the tree, represent the final class label assigned to a data point.

1. K-Nearest Neighbor

K-Nearest Neighbors (KNN) is a straightforward, yet effective machine learning algorithm used for both classification and regression tasks. It classifies or predicts the value of a data point based on the majority class or average value of its k nearest neighbors in the feature space.

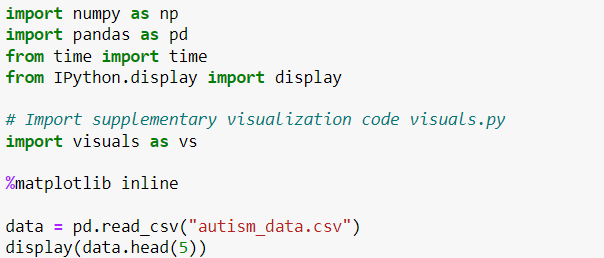
1. Support Vector Machine

At the core of SVMs is the concept of finding the optimal hyperplane that separates data points into different classes. This hyperplane is determined by a subset of data points known as support vectors, which are the data points closest to the decision boundary. The goal is to find the hyperplane that maximizes the margin between the two classes.

1. Logistic Regression

It models the probability of a binary outcome (e.g., 0 or 1, yes or no) based on one or more predictor variables. Unlike linear regression, which predicts a continuous numerical value, logistic regression predicts the probability of a specific outcome. The output of a logistic regression model is a probability value between 0 and 1, representing the likelihood of a data point belonging to a specific class.

**Libraries**



**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer code

Description automatically generated**

**Feature set Exploration**

**A screenshot of a computer

Description automatically generated**

**A screen shot of a number

Description automatically generated**

**Preparing the Data**

**A white screen with a black border

Description automatically generated with medium confidence**

**Dropping Missing Values**

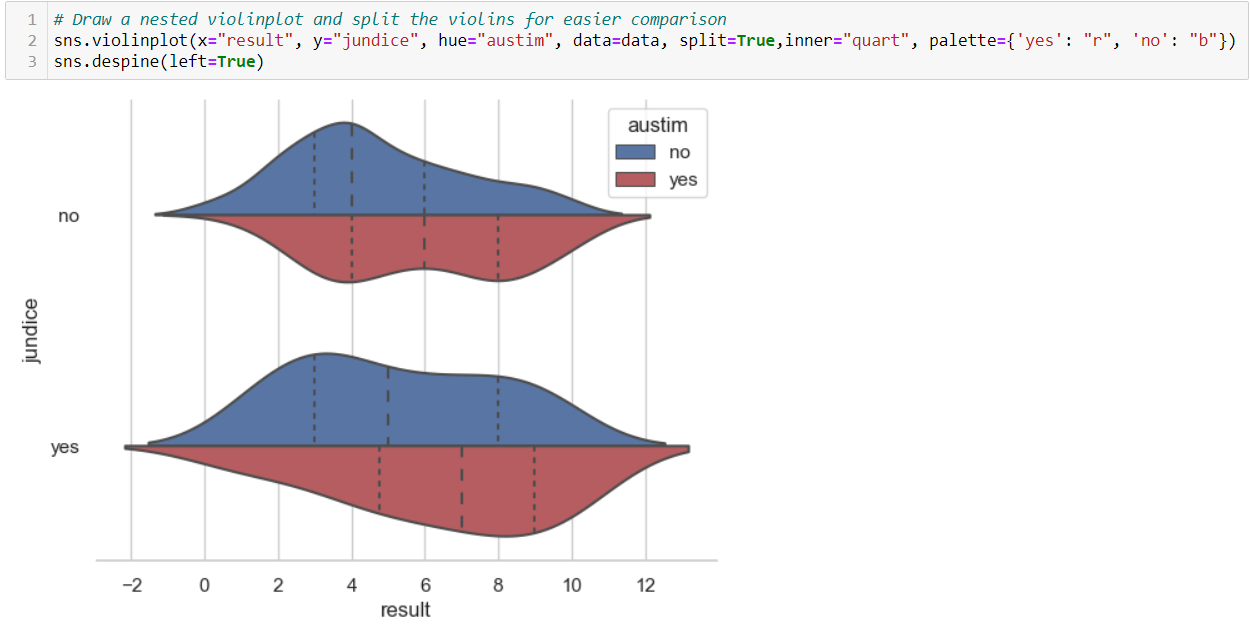
**A screenshot of a computer screen

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

**Visualizations with Seaborn**

****

**A screenshot of a computer screen

Description automatically generated**

**A screenshot of a graph

Description automatically generated**

**A screenshot of a graph

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**Data Preprocessing**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**Encode all classes data to numerical data**

**A screenshot of a computer screen

Description automatically generated**

**A white rectangular object with a black border

Description automatically generated**

**A graph of a bar graph

Description automatically generated**

**Splitting Data**

**A computer code with many letters

Description automatically generated with medium confidence**

**Models**

1. **Decision Making Tree**

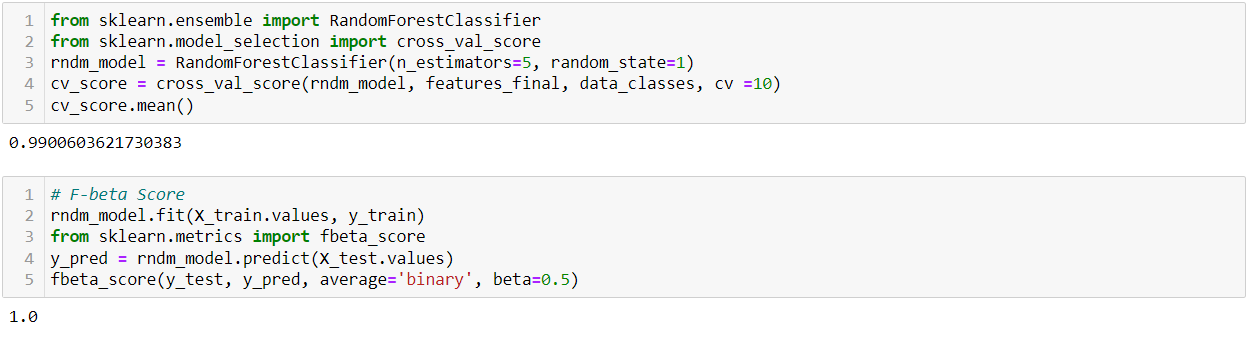
A screenshot of a computer code

Description automatically generated

**A screenshot of a computer

Description automatically generated**

1. **Random Forest Classifier**

****

1. **A screen shot of a computer code

   Description automatically generatedSupport Vector Machine**

1. **K-Nearest-Neighbors (KNN)**

**A screenshot of a computer code

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

1. **Naive Bayes**

**A screenshot of a computer

Description automatically generated**

1. **Logistic Regression**

**A screenshot of a computer program

Description automatically generated**

**Model Tuning**

**A computer screen shot of a computer code

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**CONCLUSION**

In conclusion, the application of machine learning in autism detection represents a significant advancement in the field of medical diagnostics. By leveraging sophisticated algorithms and vast datasets, machine learning models can identify subtle patterns and indicators of autism spectrum disorder (ASD) that might be missed by traditional diagnostic methods. This not only enhances the accuracy and efficiency of early detection but also makes the diagnostic process more accessible and less burdensome for patients and healthcare providers.

The integration of machine learning techniques, such as Support Vector Machines (SVM), Random Forest Classifiers (RFC), and Convolutional Neural Networks (CNN), into autism detection systems holds great promise for the future. These technologies can process diverse data inputs, from behavioral assessments to genetic information, providing a comprehensive approach to diagnosing ASD. As research and technology continue to evolve, machine learning is poised to play an increasingly vital role in improving the lives of individuals with autism through earlier and more precise diagnosis.