Application of Machine Learning in Identification of Best Teaching Method for Children with Autism Spectrum Disorder

by

Zarin Tassnim Zoana
18101677
Nasrin Akter
18101660
Mahmudul Wahed Shafeen
18101549

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

Department of Computer Science and Engineering
Brac University
May 2022

Declaration

It is hereby declared that

1. The thesis submitted is our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Zarin Tassnim Zoana

Zarin Tassnim Zoana

18101677

Nasrin Akter

Nasrin Akter

18101660

MWShafeen

Mahmudul Wahed Shafeen

18101549

Approval

The thesis titled "Application of Machine Learning in Identification of Best Teaching Method for Children with Autism Spectrum Disorder" submitted by

1. Zarin Tassnim Zoana (18101677)
2. Nasrin Akter (18101660)
3. Mahmudul Wahed Shafeen (18101549)

Of Spring, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science and Engineering on May 24, 2022.

Examining Committee:

Supervisor (Member):

Tanvir Rahman

Tanvir Rahman

Lecturer
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Md. Golam Rabiul Alam, PhD

Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD

Professor and Chairperson (CSE)
Department of Computer Science and Engineering
Brac University

Ethics Statement

Our personal findings have brought us to the conclusion that all the claims in our research are true. All of the papers and archives that were used in this paper were appropriately cited and referenced. This work or any part of it has never been submitted to another college or academic institution for the purpose of acquiring a degree.

Abstract

A good teaching method is incomprehensible for an autistic child. The autism spectrum disorder is a very diverse phenomenon. It is said that no two autistic children are the same. So, something that works for one child may not be fit for another. The same case is true for their education. Different children need to be approached with different teaching methods. But it is quite hard to identify the appropriate teaching method. As the term itself explains, the autistic spectrum disorder is like a spectrum. There are multiple factors to determine the type of autism of a child. A child might even be diagnosed with autism at the age of 9. Such a varied group of children of different ages, but specialized educational institutions still tend to them more or less the same way. This is where machine learning techniques can be applied to find a better way to identify a suitable teaching method for each of them. This paper first summarizes the research on previous researches on special educational needs of autistic children with a focus on the different behaviors that are involved in communicating with their teachers, peers and understanding capabilities that are most generally absent in children on the spectrum. The paper ends with a conclusion that uses Machine Learning algorithm in comparing different autistic traits with some suitable teaching methods, by analyzing their physical, verbal and behavioral performance. In this way, the proper teaching method can be suggested much more precisely compared to a diagnosis result. As a result, more children with autistic spectrum disorder can get better education that suits their needs the best.

Keywords: Autism Spectrum Disorder; Teaching methods; Special Educational Needs; Machine Learning; Autistic children

Dedication

To our respected faculty members who have taught us over the past four years, especially our honorable supervisor, Tanvir Rahman sir, as well as, our friends and family who have supported us all the way. we would like to dedicate this thesis to all of them.

Acknowledgement

We would like to express our humble gratitude to Tanvir Rahman, Lecturer of CSE Department, Brac University for his unwavering support, guidance, and encouragement throughout the process. We would also like to express our gratitude to Brac University for providing us with the chance and support we needed to complete this research. Furthermore, IEEE Journal version of our paper has been submitted to and accepted by CISES2022 (International Conference on Computational Intelligence and Sustainable Engineering Solutions).

Table of Contents

Declaration i

Approval ii

Ethics Statement iii

Abstract iv

Dedication v

Acknowledgment vi

Table of Contents vii

List of Figures ix

List of Tables x

Nomenclature xi

1 Introduction 1
- 1.1 Autism Spectrum Disorder and Education . . . 1
- 1.2 Research Problem . . . 2
- 1.3 Research Objective . . . 3

2 Background 4
- 2.1 Literature Review . . . 4
  - 2.1.1 Training For Teachers . . . 4
  - 2.1.2 Learning Environment For Children With ASD . . . 6
  - 2.1.3 Teaching Approach . . . 7
- 2.2 Related Works . . . 8

3 Comparative Study of Machine Learning 9
- 3.1 Methodology . . . 9
  - 3.1.1 Supervised Learning . . . 9
  - 3.1.2 Naive Bayes . . . 11
  - 3.1.3 Random Forest . . . 13
  - 3.1.4 Decision Tree . . . 15
  - 3.1.5 K-Nearest Neighbour . . . 16
- 3.2 Proposed Approach . . . 17

3.3 Input Data 20
3.4 Data Pre-Processing 22
3.4.1 Label Encoding 23
3.4.2 Feature Scaling 24

4 Experimentation **25**
4.1 Machine Learning Model 25
4.1.1 Implementation 25

5 Results **26**

6 Conclusion **28**

Bibliography **34**

List of Figures

3.1 Supervised Machine Learning Working Procedure 9
3.2 Confusion Matrix Diagram 12
3.3 Random Forest Algorithm Working Procedure 13
3.4 Decision Tree Working Procedure 15
3.5 K-NN Algorithm Working Procedure 17
3.6 The Flow Chart of the Proposed Research 18
3.7 Label Encoding 23
3.8 Feature Scaling 24

6.1 Algorithms and Scores 28
6.2 Bar diagram of Accuracy, Precision, Recall, F1 Score 29

List of Tables

4.1 Preferred educations based on A1 to A10 scores 25

5.1 Accuracy Table 26
5.2 Precision Table 26
5.3 Recall Table 27
5.4 F1 Score Table 27

Nomenclature

The following rundown depicts a few images and condensing that will be subsequently utilized inside the body of the report

ABA Applied behavior Analysis

AI Artificial Intelligence

ASD Autism Spectrum Disorder

CBI Cognitive Behavioral Intervention,

EBE Evidence Based Education

$fANN$ fast Artificial Neural Network

$K - NN$ K-Nearest Neighbour

ML Machine Learning

Chapter 1

Introduction

1.1 Autism Spectrum Disorder and Education

Autism Spectrum Disorder or ASD is a disorder of neurological development which mostly occurs during the early stage of infancy, that is distinguished by a persistent deficiency in daily life interaction and communication and repetitive patterns of behavior [22]. As per the information of World Health Organization, about one in 270 people globally is estimated to have ASD [41]. And 31% of autistic child have intellectual quotient (IQ) of less than 70, that is, they have an intellectual disability [38]. That is why autistic students find it hard to obtain the same academic attainment as opposed to other students [4].
According to Universal Declaration of Human Rights' Article 26, everyone has the right to education [1] , claiming all types of students, like children with disabilities and from minority groups from all corners of society have the right to access education, and that includes children with Autism Spectrum Disorder. In the past two decades, a significant increase was noticed in research on Autism Spectrum Disorder [18] and following this, the tendency of including autistic children in the general and specialized education system has also ascended.
However, special education for students with Autism Spectrum Disorder is far from perfect. Students with ASD require a special form of education and are thus considered as students with 'special educational needs' or SEN [40]. Arranging this SEN is mostly hard not only due to the students lacking communication and interactional skills, but also due to lack of collaboration between teachers and parents, and a very limited amount of practical suggestions and proper resources to guide them [40].
Each year more research is done on Autism Spectrum Disorder and specialists are able to diagnose children with ASD much more successfully than before. In the last decade, the number of children diagnosed has increased an incredible five times [18]. Based on their diagnosis, the children are categorized into their own sections of special education. But quite often, they are rejected in general schools, and even when they get their primary education in a specialized school, later in life, they drop out during their higher studies, due to anxiety and poor academic achievement.
Most researches on Autism Spectrum Disorder are done using large-scale statistics. Since diagnosis is done based on statistics and not persistent data, there might be some inaccuracies. And thus, identifying their special educational needs is not so accurate as well. But this can be done more efficiently using Machine Learning algorithms that can solve large amounts of variables more precisely.

1.2 Research Problem

This research focuses on betterment of special educational needs of children in the spectrum of autism through Machine Learning. However, such an approach has not been taken before for any research in this field. Machine Learning refers to a type of Artificial Intelligence (AI) and computational science which deals with data as well as algorithms and imitates how humans learn and gradually improves its accuracy [37]. In this research, data taken from schools for children with special needs will be taken and using Machine Learning algorithms, an algorithm will be developed that will store all this information. It will give an output of one or a combination of special education method(s) for any children after comparing his/her given input with all the other stored information. However, there are certain challenges for which Machine Learning has not been used before to develop any strategy to help children with special education needs because Machine Learning requires a large dataset to work efficiently. But a lot of people tend to hide and deny their or their child's disorder. This is termed as camouflaging autism. That is why datasets of autistic children are not a hundred percent accurate. Besides, concern about privacy of an individual's information is also a key factor behind not trusting this kind of approach. From a survey conducted by Individuals with Disabilities Education Act (IDEA), 7.3 million students aged between 3 to 21, received special education services which is 14 percent of all students in public schools, and only 11% of them have been found to have autism spectrum disorder [20]. That is actually 1.54 of total population. That is only half of the total population thought to be diagnosed with ASD as mentioned by WHO [41]. So, at its best, this process can only use data from half of the real number of people with ASD. Moreover, machine learning can be biased and discriminating when the available statistics or data being used are slightly undistributed. And since autism spectrum disorder is heterogeneous, if data from one small group of children is missing in the predefined dataset, that could hugely impact the results provided by the algorithm. This type of bias is called the Missing-the-Forest-for-Trees bias, where the algorithm might have blind spots and these blind spots can create bias in the outputs [46]

Another issue regarding special teaching methods for autistic children is unawareness of parents, local guardians and people in general, so they are not easily intrigued in such approaches. In a researched carried out in Saudi Arabia, it was found that, among a total of 259 individuals with (n=47%) less than 30 years old, (n=57%) females, (n=60%) married, (n=68%) educated, (n=54%) employed participants, 12% of them had no idea about autism spectrum disorder and 41% agreed to have weak knowledge about ASD [28]. That is why, the result needs to be presented in a more user-friendly way, so that the diagnosed children and their parents and all people in general can understand the whole process of determining the required special education method.

Therefore, our research is trying to answer the following question: **How efficient is machine learning in determining the best teaching method for children with autism spectrum disorder?**

1.3 Research Objective

Children having autism spectrum disorder may find it hard to understand and communicate their needs to teachers and peers. They can have trouble understanding some directions and instructions in the classroom. It is also hard for them to interpret subtle vocal and facial hints from their teachers. So, knowing how to teach these children and which strategies to use, is utmost important. This research seeks to provide every child with ASD the best education possible, computed by machine learning techniques.

The objectives of this research are:

1. To recognize the teaching barriers for children with ASD.
2. To know how the different teaching strategies work.
3. To understand the machine learning techniques that we can use to solve the problem.
4. To develop a model for detecting one or a combination of more than one teaching method(s) based on algorithms and apply it to find the best possible solution. To analyze the model.

Chapter 2

Background

2.1 Literature Review

An astonishing amount of researches on various types of education have been done, that could be considered best for students with autism spectrum disorder, and also how to implement them. In all these researches some factors can be found that are common in terms of implementing different methods of special education. Three such factors are very important in deciding the perfect method of special education. These three factors have been discussed in the following.

2.1.1 Training For Teachers

Report shows that students with ASD are prone to have higher behavior difficulties compared to the students who have intellectual incapacity and with ordinary improvement issues [33]. Almost 44% of students with ASD fall into the category of having higher behavior problems [26]. So, these students need special treatment and experienced teachers who have skill and experience dealing with them. Children with ASD are different from each other. One education approach that works for some students is not applicable for others [6] [11]. But statistics show that only 5% teachers use special methods and approaches specified for ASD students [19]. This shortage is caused by appointing less experienced teachers and lack of training for them to treat children with ASD. First of all, the bond of students and teachers form the early school is shown to be a greater success in the long term and developing the children [14] [25] [5] Positive and supportive behavior towards students helps them to build a strong social and communication skill. But in reality, students with ASD have poor relationships and less closeness [25] [27]. Teachers should be trained in such a way so that they can treat the children with patience and positive attitude. On the other hand, the relationship between teacher and parents is important because in this way the teacher will be able to know the needs and lacking of his student and also vice versa. Secondly, not enough training for teachers is another important issue regarding the lack of experienced teacher for students with ASD. Trainings that are being offered for general education of teacher do not emphasis based on any particular social language and educational needs for
students with ASD [15]. Rather it emphasis more on instructional strategy and accommodations. This does not provide a wide knowledge and skill. So, in the result neither they are prepared for including special students in the general education

classroom nor to assist them with individualized educational needs. Again, the report shows that only 5% of teachers use the best practice for teaching students with ASD and although some teachers apply them, the approaches they use do not differ from those who do not use the best practice [9]. Teachers should be encouraged to attend the training and the quality of these training needs some improvement. Pre-service and in-service training should include the best practice so that they have the appropriate knowledge for teaching students with ASD. Parents-teachers meeting should be required for better development of the children not only at school but also at home.

2.1.2 Learning Environment For Children With ASD

Flexible classroom management is considered as a vital issue for children with autism. As we know every child's behavioral pattern is different so they will need different equipment from which they can adapt with the learning process. Many environmental aspects such as passageways, furniture, the sights, sounds, decoration, seats, classroom and school size can have a favorable or negative impact on the children's concentration. Schools should be compatible with the demands of all students, regardless of their ability and circumstances. However, there are no explicit rules for using the classroom atmosphere as a learning environment to meet the requirements of those children who have difficulty in concentrating. The ability of autistic children to concentrate is affected by the elements of their learning environment. Thus, these must be taken into account while allocating them in the school building [13]. Along with that, the classroom should be equipped with enough modern tools. Learning relies on visualization. An immersive room decked with a 320-degree curved screen and a high end five panel projection system, virtual and motion-based interaction can be used to engage children with autism [21]. There is a significant drawback in this method. In this procedure, the expanse will increase but it will be effective for teaching and building specialized classroom.

Developing an adaptive curriculum can help creating an effective learning environment for the children with ASD. Here, didactic strategies can be used for developing the curriculum. The didactic approach implies the recognition and categorization of two components. The first one is the task of teaching and the second one is the learning situation. Both of them are aimed at establishing the best possible environment for children to attain a given task [23]. For the learning environment, a variety of constructive didactic methodologies can be implemented. A major goal of the didactic approach is to encourage children with ASD to learn more deeply about a vriety of relevant topics [10].

2.1.3 Teaching Approach

Since students with Autism Spectrum Disorder face problems in communicating and interacting with their teachers, the teachers have to make sure they are taking the proper approach that will create a path of communication as well as be effectively able to teach the students the particular topic or skill to be taught. Many researchers have come up with different teaching strategies for students with ASD. One of these strategies is called the Evidence Based Education (EBE). It refers to education practices that are based on the most recent scientific evidence, instead of perspectives, beliefs, traditions or any other factors [12]. EBE is considered the best strategy for children with ASD because Evidence-based education has two operating levels: (1) It utilizes available evidence from researches all over the world and literatures on teaching method for autistic children and then (2) creates reasonable evidence where there is none, or in case it is doubtful, ambiguous or weak [3]. ABA, or Applied behavior Analysis, is another essential strategy for teaching students with ASD. This refers to a systematic approach in understanding a behavior and the use of that information to improve socially relevant behaviors, particularly in the field of developmental disorders [29]. With the help of the techniques of Applied Behavior Analysis, researchers have drawn 27 evidence- based practices to be used as teaching methods for autistic children. These are: (1) , Cognitive Behavioral Intervention, or CBI, (2) Antecedent-Based Intervention (3) Discrete Trial Teaching, (4) Differential Reinforcement, (5) Extinction, (6) Exercise, (7) Functional Communication Training,, (8) Functional Behavior Assessment (9) Naturalistic Intervention, (10) Modelling,
(11) Peer-Mediated Instruction and Intervention (12) Parent- Implemented Intervention, , (13) Picture Exchange Communication System, or PECS, (14) Pivotal Response Training (15) Reinforcement, (16) Prompting, (17) Scripting (18) Response Interruption, , (19) Social Narratives, (20) Social Skills Training,, (21) Structured Play Groups, (22) Self-Management (23) Technology-Aided Instruction and Intervention, (24) Task Analysis, (25) Visual Supports Time Delay, , and (27) Video Modeling [8].
Another excellent method to ensure educating students with autism spectrum disorder is the Errorless Teaching Method. It is called errorless because this instructional strategy ensures that the student always responds correctly. Errorless teaching was developed by assuming that children acquire concepts and skills in the effective way through positive responses and exercises rather than making mistakes [2]. This method works in three steps. Firstly, the subject to be taught should be determined and presented to a student based on his capacity. Then, the teacher should divide the subject into small parts and teach only one at a time and develop skill analysis only when necessary. Finally, different teaching strategies (the 27 methods mentioned earlier) should be applied to facilitate student learning [43]. Errorless teaching is considered highly effective due to its advantages, where the teacher helps the students respond correctly to almost any attempt, and even if they make mistakes they can learn from it [43]. However, the most important thing of all is the collaboration between teachers and students and their parents. From that collaboration, the institution has to create a clear school policy [40] and develop an adapting curriculum with regard to students with Autism Spectrum Disorder.

2.2 Related Works

This section will evaluate previous relevant machine learning work in the context of determining the optimum teaching method for autistic children. We examine the different strategies utilized to accomplish the major conclusions and explain how matching learning has its own set of challenges and limitations in this regard. Since the past decade machine learning approaches have been very convenient to diagnose ASD in children. For instance, Virginia Tech Centre along with Virginia Tech Institute for Society, Culture and Environment (ISCE), did a research where they applied fast Artificial Neural Network (fANN) technique using data from 14,995 infants (16-30 months, 46.51% male) taking twenty inputs and produce an assertive or negative output if the kid has ASD. The sample was then clustered into groups according to their race, sex, and maternal education (i.e., mother has completed Associate Degree or not). The results yielded 99.72% accurate, with 99.92% accuracy for white children and 99.79% for black children. For the boys the results were 99.64% correct, while it was 99.95% for the girls. In case of maternal education, the results were 99.75% correct for mothers having completed the degree and 99.70% for mothers who have not. [32] However, very recently, focus has been shifted to use of technology in the educational needs of special children. Department of Computer Science and Institute of Education, University College London used machine learning techniques to figure out how students respond to various forms of communication used by their teachers inside controlled classroom conditions considering their specific attributes. [44] There is a lot of space for further research in this field of machine learning approaches to find the proper teaching method for children with ASD.

Chapter 3

Comparative Study of Machine Learning

3.1 Methodology

3.1.1 Supervised Learning

Supervised learning is one of the three types of machine learning, that is a sub-category of artificial intelligence. This type of machine learning is known for using labelled datasets to determine patterns, create classification of the data, and calculate estimated results as precisely as possible. It takes a set of data to learn patterns and educate particular models and then yields the desired output. The model later incorporates more inputs to the dataset and compares outputs, which allows it to learn over time and become more accurate in giving outputs. The model determines its accuracy by implementing the loss function, updating itself until the error is adequately reduced to the point where it is negligible [49].
The working procedure of Supervised machine learning is illustrated in the following diagram:

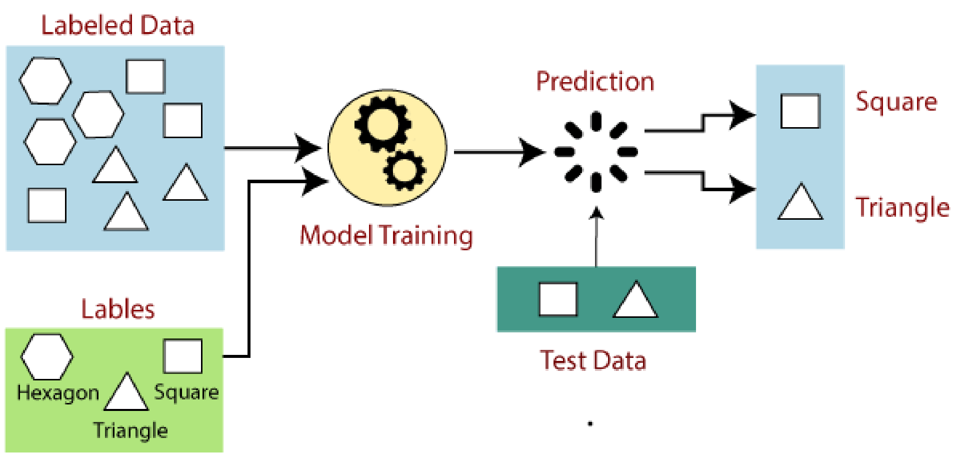


Figure 3.1: Supervised Machine Learning Working Procedure

We decided to use a Supervised learning algorithm as it permits collecting information and produces output from past encounters. Makes a difference to optimize execution criteria with the assistance of past experience. Different algorithms and computation methods are utilized in supervised machine learning processes. We have decided to work on these specific four algorithms for our research.

3.1.2 Naive Bayes

Naive Bayes classifier is known as probabilistic classifier, that is based on the Bayes hypothesis with presumption of independence among predictors. Bayes hypothesis depicts the probability of an event to occur by comparing it to the previous records of the event's probability. It is a classification approach that uses the theory of class conditional independence from the Bayes Hypothesis. This implies that the existence of one feature does not impact the existence of another within the probability of a given result, and each predictor has a rise to impact on that result [34].
Given a problem instance, represented by a vector x= (, to be classified using Naïve Bayes; the model assumes that one feature in a class is unrelated to any other feature in that same class.

$$P(c_k|x_1, x_2 \ldots \ldots \ldots x_n)$$

$$Posterior = \frac{Perior * Probabilty}{Proof}$$

The denominator is essentially a constant in this case.
Steps:
1) The dataset is converted into a frequency table.
2) After determining the probability, a table is created.
3) Using the Naïve Bayesian equation, the posterior probability is calculated.
4) The class containing the highest posterior probability forms the predicted outcome [50].

Confusion Matrix

Confusion Matrix represents the performance estimation of two or more classes of outputs from classification done with a machine learning model. It is a matrix used to portray the execution of a classification model by comparing a set of test data, whose true values are known, and a similar set of predicted data. Confusion matrix can be very useful for measuring Review, Accuracy, Specificity, Precision, and most imperatively AUC-ROC bends [36].

$$Classification\ Accuracy = \frac{Correct\ Predictions}{Total\ Number\ of\ Predictions}$$

It is also possible to roll the classification accuracy into an error rate by simply calculating the inverse value, such as:

$$Error\ Rate = \left(1 - \left(\frac{Correct\ Prediction}{Total\ Number\ of\ Prediction}\right)\right) * 100$$

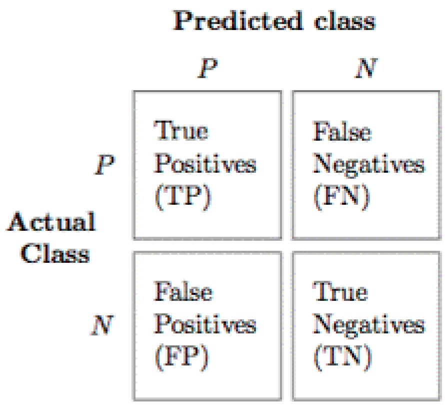


Figure 3.2: Confusion Matrix Diagram

3.1.3 Random Forest

Random Forest is a popular supervised machine learning algorithm. It can be used as a Classification model as well as a Regression model. It is based on the notion of supervised learning, which combines multiple classifiers and solves complex problems and moves forward to execute the model. It is a classifier that combines a bunch of decision trees on different parts of a particular dataset and uses the normal to improve the dataset's predictive precision. Rather than relying on a single decision tree, the random forest takes the output from each decision tree and predicts the final output based on the majority of predictions. The larger number of trees within the forest, the more precise it is and avoids any complication of overfitting [35]. The working procedure of the Random Forest algorithm is shown below:

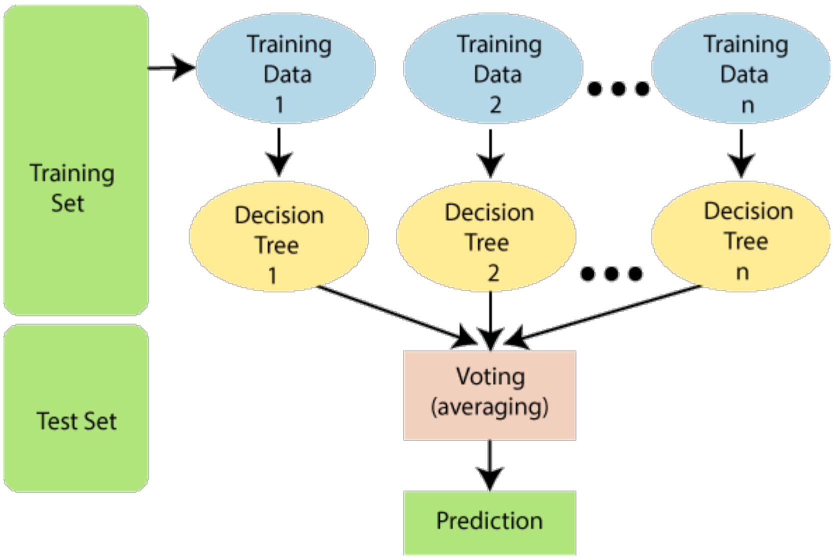


Figure 3.3: Random Forest Algorithm Working Procedure

To solve regression problems, the Random Forest algorithm uses the mean squared error (MSE) to calculate how the data branches from each node.

$$MSE = \frac{1}{N} - \sum_{i=1}^{N} (f_i - y_i)^2$$

Where N is the number of data points
fi is the value returned by the model and
yi is the actual value for data point i.

This algorithm determines the distance between each node's expected actual value and the predicted actual value, assisting in determining which branch is the better output for the random forest. yi is the value of the testing data at a particular node and fi is the value returned by the decision tree. Again, when running Random Forests on the classification data, the Gini index or this formula is applied:

$$G_i n_i = 1 - \sum_{i=1}^{C} (p_i)^2$$

The above formula uses class and probability to calculate the Gini of each branch on a node, discovering which of the branches is more likely to occur. pi is the relative frequency of the class that is being observed in the dataset and c is the total number of classes. In Random Forest, Entropy can also be used to picture how nodes branch in a decision tree.

$$Entropy = \sum_{i=1}^{C} -p_i * log_2(p_i)$$

Entropy makes a decision on how the node should branch based on the probability of a certain outcome to occur.

3.1.4 Decision Tree

Decision tree is known as an effective algorithm in the case of prediction as well as classification. It can be used in the supervised learning method. A Decision tree looks like a flowchart type of tree structure and on each inner hub indicates a test on an attribute. In the decision tree algorithm, each branch represents an output of the given input-based testing, and each leaf node, also called terminal node, holds a class label. In a Decision tree, there are two nodes that are being used, first one is Decision Node and the other one is Leaf Node. Mainly, Decision nodes are utilized to make any decision. Decision tree algorithm has numerous branches, while its Leaf nodes are the output of those decisions and do not contain any further branches [47]. On the premise of features of the given dataset, here almost all decisions are performed. It is a graphical representation for getting all the conceivable outputs to a problem based on given conditions. There is a specific reason to call it a decision tree because it begins with the root node that extends on further branches comparable to a tree and builds a tree-like structure. A decision tree basically inquires, and according to the reply (Yes/No), it assists parting the tree into subtrees.

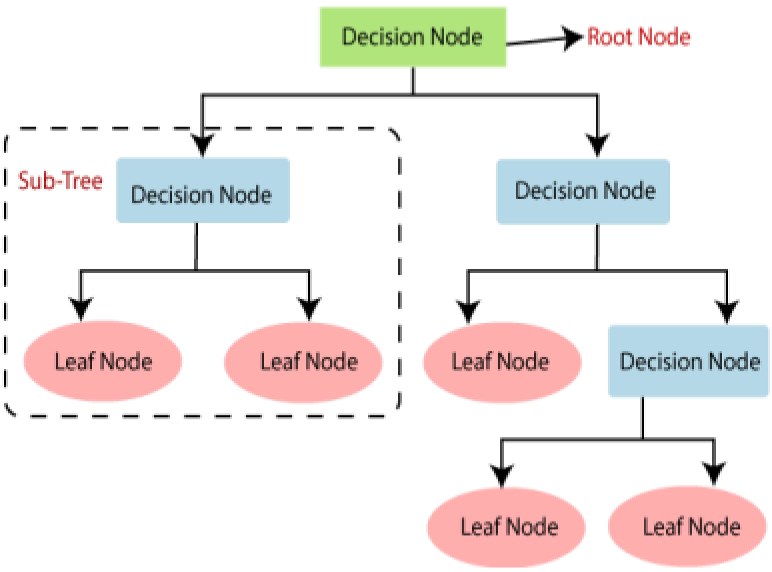
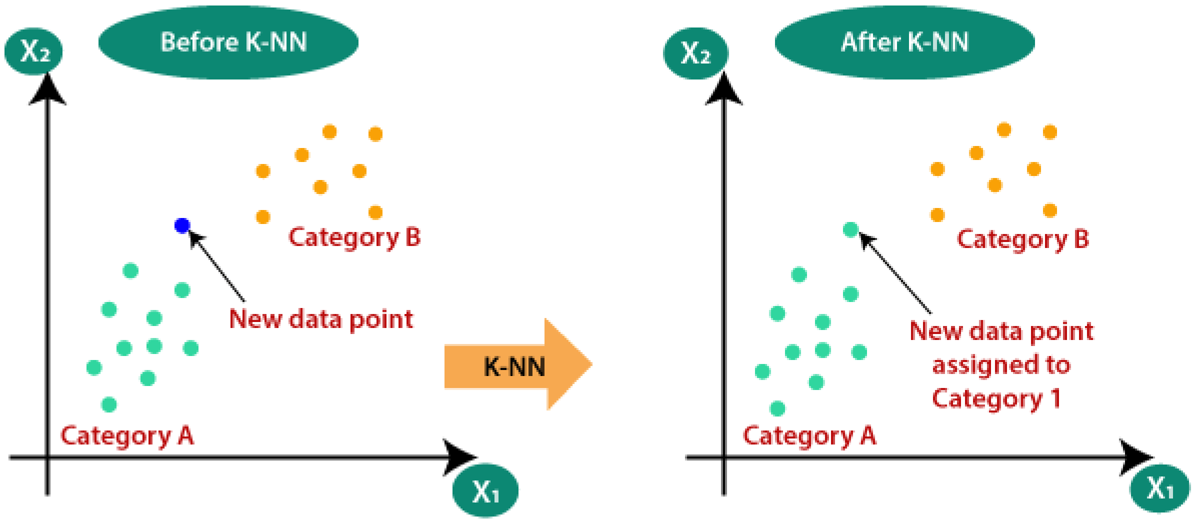


Figure 3.4: Decision Tree Working Procedure

In Machine learning, choosing the finest algorithm for any specific dataset is quite an important point to notice while creating a machine learning model. Underneath are the two reasons for utilizing the Decision tree:
1) Decision Tree is simple to understand because it more often than not mimics humans if considering the capacity while making a decision.
2) It appears as a tree-like structure so the logic behind the decision tree can be easily understood.

3.1.5 K-Nearest Neighbour

Another widely used supervised machine learning approach is K-Nearest Neighbor. The K-Nearest Neighbor technique assumes that the unused data and accessible cases are close in proximity, and places the new data in the category that is most similar to the available categories. The K-NN algorithm saves all available data and categorizes unused data points depending on their proximity. This means that as new information appears, it may be quickly sorted into a suitable category using the K-NN algorithm. The K-NN algorithm can be used for both regression and classification, but it is more commonly used for classification tasks. It is also known as a lazy learner algorithm since it does not learn from the training set right away; instead, it saves the dataset and performs an action on it when it comes time to classify it. At the training stage, the K-NN algorithm simply saves the dataset, and when it receives new data, it classifies it into a category that is significantly more comparable to the new data [48]. Assuming there are two categories, Category A and Category B, and we have an unused data point x1, which of these categories will this data point go into? A K-NN algorithm is required to solve this type of problem. We can effectively recognize the category or class of a certain dataset with the help of K-NN.



Steps of K-Nearest Neighbor (KNN) Algorithm:
Step-1: Decide on the number of the neighbors, K.
Step-2: Determine the Euclidean distance between K neighbors.
Step-3: Using the estimated Euclidean distance, select the K closest neighbors.
Step-4: Count the number of data focuses in each category among these k neighbors.
Step-5: Assign the new data focuses to the category with the highest number of neighbors. Step-6: The model is ready to use.

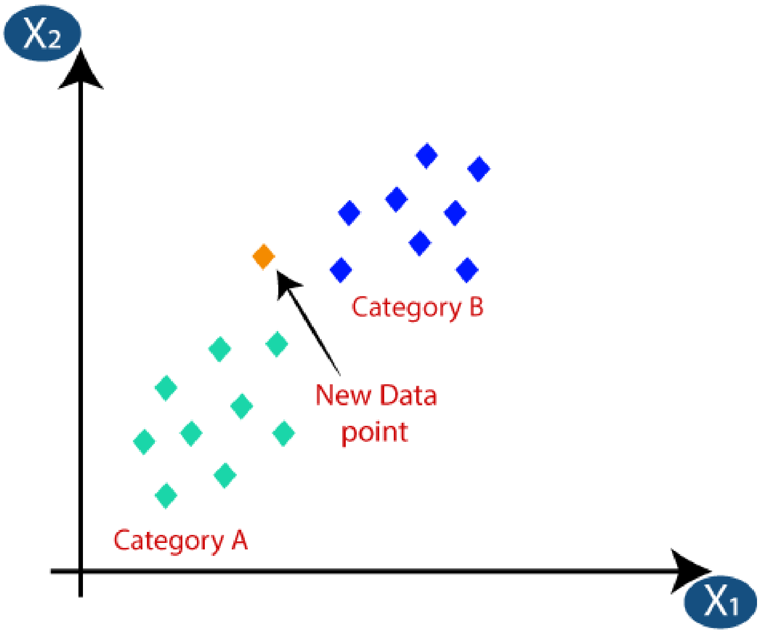


Figure 3.5: K-NN Algorithm Working Procedure

3.2 Proposed Approach

The goal of the proposed model is to apply machine learning to determine the best teaching method for children with ASD. To accomplish said goal, the model needs to design a process that accepts data as an input, process it, apply the machine learning algorithms and find the best fit. The figure below provides an effective view of the model design. First, we have to collect data relevant to the research. The input data preprocessing stage is concerned with dropping unnecessary data and converting the datatypes of necessary data to make it easy to process in the model. The preprocessed input data goes into the train and test splitting stage and builds two segments one used to train the model and the other used to compare with the predicted outputs. After that, selected machine learning algorithms are used and run on the preprocessed input data for selecting the best fit algorithm and apply them in finding the appropriate teaching method for children with ASD.

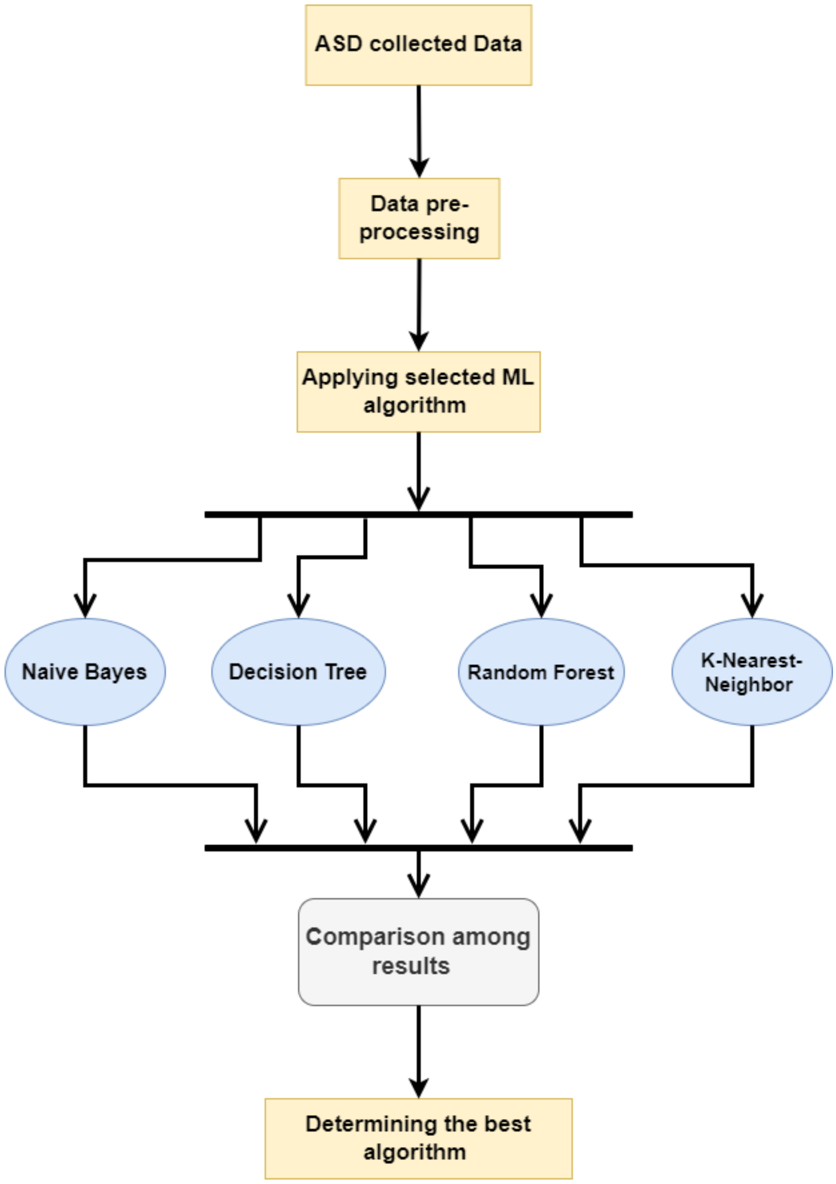


Figure 3.6: The Flow Chart of the Proposed Research

Among the many education methods discussed in the Teaching Approach section of the Literature Review, six teaching methods have been selected for the project. The first one is Technology-aided instruction and intervention, which is a learning system that uses technology as its central feature. This is designed for children who need some extra time to go over an exercise and adjust to their own pace of learning [30]. As a result, teachers and other students needing less support can continue ahead. Secondly, Antecedent based intervention is another evidence-based practice that identifies what causes interference to a child and modifies the environment to remove said interfering behavior of the child so that he/she can regain focus on the exercise [17]. The third teacher method used in the project is Pivotal response training. This is a therapy type training system that increases a child's motivation to learn, start communication with someone, and monitor their own behaviors [24]. On the fourth spot, we have Peer-mediated instruction and intervention. This method involves another child without disabilities to take on a role in the teaching beside the teacher or therapist [16]. The peer not only plays the role of a tutor, but also teaches critical social skills along the way. Next is the Picture Exchange Communication method, which encourages verbally challenged children to use visual symbols to communicate with parents, teachers and peers. It is a kind of complementary and alternative system that is used for intentional and functional communication to tell what they want or need [7]. Lastly, we have selected Task Analysis, which breaks down complex tasks into sequential smaller steps. This teaching method is applicable for individuals who find even simple tasks to be difficult and give up [39].

3.3 Input Data

As for input data to determine the best teaching method for children with autism, it was quite difficult to obtain data, since such data are confidential, and very few researches have been done on this. We are using a supervised Machine Learning method for our research and so, we have searched for labeled data. Since there was no dataset large enough to be used for Machine Learning, we have used a total of 4 datasets for our research and run our algorithms on a merged dataset with the same attributes.
We have used the dataset named 'Autistic Spectrum Disorder Screening Data for Toddlers' documented by Dr. Fadi Fayez Thabtah, Principal Lecturer of Digital Technology department of Manukau Institute of Technology in New Zealand [31].
Download link : https://www.kaggle.com/datasets/fabdelja/autism-screening-for-toddlers?select=Toddler%2Bdata%2Bdescription.docx

Second dataset that we used was named Behavior Analysis of Autism.
Download Link : https://www.kaggle.com/datasets/iashiqul/behavior-analysis-of-autism

The other two data sets we have used the data set named Autism screening child consists of two version. It has 2249 instances, 18 attributes.
Download Link : https://www.kaggle.com/datasets/basmarg/autism-screening-child-two-version

Data sets regarding clinical or screening autism spectrum disorder are presently very limited and also generic. That is why we used this merged dataset because it is based on influential features that can be useful for further research in improving the classification process and determining ASD cases. The dataset contains categorical, binary and continuous data. The information is focused on medical, health and social science areas. The data set holds 3043 samples of different children with autism spectrum disorder. The dataset holds 3043 samples of different children with autism spectrum disorder. The dataset contains 12 integer type data including case number, closed questions whose answers are represented by binary 1's and 0's and Qchat- 10 score. The other 7 data are object type. The total size of our merged dataset is 451.8 kB. The dataset contains 10 closed questions whose answers are represented by binary 1's and 0's. Answers of the later 9 questions are represented by strings and numbers.

Table 1: Details of variables mapping to the Q-Chat-10 screening methods

Variable in Dataset	Corresponding Q-chat-10-Toddler Features
A1	Does your child look at you when you call his/her name?
A2	How easy is it for you to get eye contact with your child?
A3	Does your child point to indicate that s/he wants something? (e.g. a toy that is out of reach)
A4	Does your child point to share interest with you? (e.g. poin9ng at an interes9ng sight)
A5	Does your child pretend? (e.g. care for dolls, talk on a toy phone)
A6	Does your child follow where you're looking?
A7	If you or someone else in the family is visibly upset, does your child show signs of wan9ng to comfort them? (e.g. stroking hair, hugging them)
A8	Would you describe your child's first words as:
A9	Does your child use simple gestures? (e.g. wave goodbye)
A10	Does your child stare at nothing with no apparent purpose?

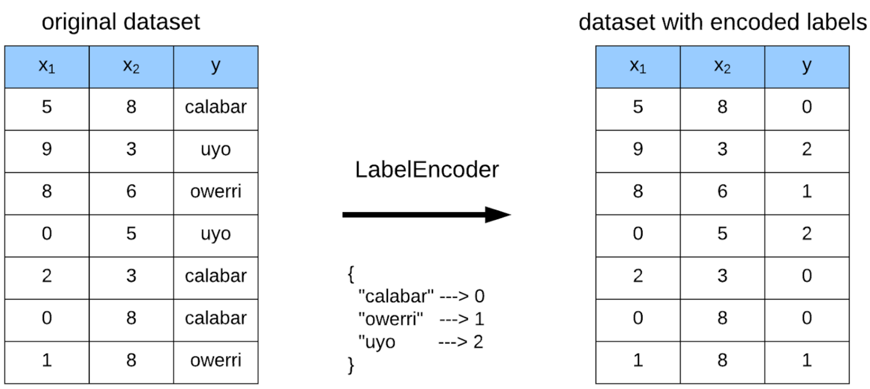
We have used a merged dataset that consists of 4 individual datasets that has 3043 instances in total. We decided to use this dataset because the problem with scarcity of data is a crucial drawback, as a good amount of data is integral to any project done with machine learning algorithm. If a dataset is inadequate, it might as well be the agent of poor performances in the project. Very often, drawbacks like these are the main reason why significant machine learning projects remain unaccomplished. Most supervised learning algorithms are strongly reliant on the amount of training data provided. It can be challenging to create large enough training datasets in many circumstances. With smaller datasets algorithms tend to learn the detail of the noises in the training data to such extent, that the model's performance on newer data can have a negative impact. In general, the smaller the dataset, the better it is to use the simpler the machine learning algorithm. Small data requires low-complexity models in machine learning in order to prevent overfitting the model to the data. The Naive Bayes for example, is one of the most basic algorithms, and as a result, its performance on learning from comparatively minor dataset can indeed be remarkably well. Furthermore, other simpler algorithms such as decision tree, random forest and K-NN can also learn really well from small datasets. These algorithms are essentially better than more complicated algorithms in case of learning from smaller datasets, as they actually try less to learn from exceptional cases or noises.

3.4 Data Pre-Processing

Now that we have the merged dataset, we initially tried to run four Machine Learning algorithms through the dataset without any kind of pre-processing techniques. That means, we only got to use the integer data types for training the algorithms. These were the A1-A10 answers and Qchat-10 score. However, three out of four algorithms gave 100 percent accuracy and the other one gave 99.8 percent accuracy. This case can be interpreted as such that there was no point in using Machine Learning and this could be done only with a bunch of if-else conditions. That is when we decided to include more features of object datatype. So at first, we used the Label Encoding pre-processing method to convert the object datatypes into integer datatype so that the algorithms can read those features as well. By doing so, we trained the algorithms with 16 features instead of 11 features done previously. With more input data, a machine learning algorithm can learn patterns more efficiently. However, due to Label Encoding, some data became unbalanced, so after splitting the training and testing sets, we used Feature Scaling pre-processing on the training set to achieve a higher accuracy rate.
However, due to Label Encoding, some data became unbalanced, so after splitting the training and testing sets, we used Feature Scaling pre-processing on the training set to achieve a higher accuracy rate.

3.4.1 Label Encoding

Most datasets consist of multiple labels in one or more of its columns, which are string datatype, for example: words or big numbers. These words or numbers are convenient for humans to understand the dataset, but it is complex and time consuming for the algorithms to read.
This is where Label Encoding is necessary. This preprocessing method converts categorical labels into numeric forms to make it machine-readable. It is a very simple and straightforward process that changes all the labels in a column into integers from "0" to "n-1", n being the number of unique datas in that column. The numbers dedicated to the labels do not hold any priority relation among the labels.



original dataset

x_1	x_2	y
5	8	calabar
9	3	uyo
8	6	owerri
0	5	uyo
2	3	calabar
0	8	calabar
1	8	owerri

dataset with encoded labels

x_1	x_2	y
5	8	0
9	3	2
8	6	1
0	5	2
2	3	0
0	8	0
1	8	1

Figure 3.7: Label Encoding

3.4.2 Feature Scaling

Most datasets also consist of independent features in a single column, which tend to have a higher range between the maximum and minimum data, or have different units (like metres and miles). It is necessary to handle this huge variability.
Feature Scaling is a pre-processing technique that standardizes or normalizes these features to bring them to a fixed range, or similar units.

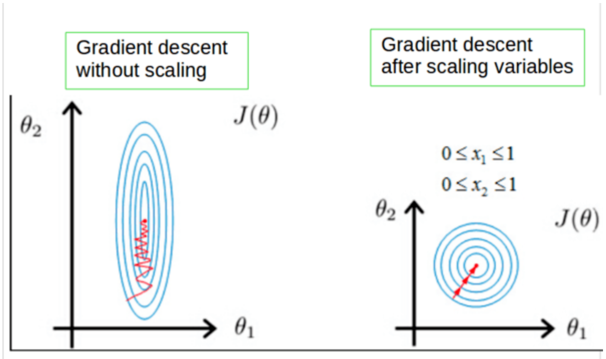


Figure 3.8: Feature Scaling

Chapter 4

Experimentation

4.1 Machine Learning Model

The objective of our project is to use a Machine Learning Algorithm that will learn from patterns in autistic traits, sex, and other factors responsible for autism and then predict a preferred educational method for an individual. Due to its access to many built-in libraries such as numpy, pandas, matplotlib, sklearn, the most preferred language Python has been used in the project. The project is done in Google Colaboratory. The datasets used for the project have been stored in Google Drive in the same folder as the Colab Notebooks.

4.1.1 Implementation

The project file starts by importing the necessary libraries and frameworks. At the next step Google Drive is mounted with Colab and the dataset is read by the project file using pandas -library, which is the best way to handle large data frames in a Python program. For simplicity and efficiency, we preprocess the dataset with the help of two methods. We used Label Encoding to include more machine-readable features to train the algorithms. The binary data A1 to A10 and Q-chat-10 Score are kept as is, while the rest of the columns that had object datatype are converted to integer datatype and added as new columns. We also used Feature Scaling on the training data to balance out irregularly encoded datas. Another new column namely 'Preferred education' is added using the numpy library. This column consists of integers 1 to 6 to represent the six special teaching methods, and 0 to represent that no special education is required. These six teaching methods are selected based on some conditions set upon the binary datatype autism traits A1-A10, as followed. Table 3: Preferred educations based on A1 to A10 scores:

Table 4.1: Preferred educations based on A1 to A10 scores

Technology-aided instruction and invention	Antecedent based interven tion	Pivotal response training	Peer-mediated instruction and invention	Picture exchange communication	Task analysis
A5=1	A6=1	A1=1	A5=1	A2=1	A7=1
A9=1		A8=1	A4=1	A9=1	
A10=0			A3=1		

Chapter 5

Results

At first, the columns of dataset are divided to 2 sets, one as independent variable x, which includes A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, Q-chat-10 Score, Age Mons enc Sex enc, Ethnicity enc Jaundice enc, and Family mem with ASD enc and Class/ASD Traits enc. The other set of dependent variables y consists of just the 'Preferred Education' column. In the next step we split both x and y part of the dataset to the training and testing set. To be more accurate, the testing part is set to be 0.05, i.e. 5% of the total data. That means that the algorithm will learn from 95% of the data. It will find patterns from the x and y of the training set and then predict y from the x of the testing set. Finally, the predicted y values are checked with y values of the testing set to calculate accuracy of the algorithms. Here are the Confusion matrices and Accuracy score of all four algorithms.

Table 5.1: Accuracy Table

Algorithms	Accuracy
Naïve Bayes	90.85%
Random Forest	98.69%
Decision Tree	98.69%
K-NN	94.12%

Table 5.2: Precision Table

Algorithms	Precision Score
Naïve Bayes	91.70%
Random Forest	99.10%
Decision Tree	97.55%
K-NN	92.64%

Table 5.3: Recall Table

Algorithms	Recall Score
Naïve Bayes	95.24%
Random Forest	97.95%
Decision Tree	98.80%
K-NN	91.88%

Table 5.4: F1 Score Table

Algorithms	F1 Score
Naïve Bayes	92.19%
Random Forest	98.48%
Decision Tree	98.08%
K-NN	92.10%

It is easily understandable that both Random Forest and Decision Tree have the highest accuracy of approximately 98.69 percent. But Random Forest has the highest precision score of 99.09 percent and Decision Tree has the highest recall score. Now the highest F1 score of Random Forest, i.e. 98.47 percent breaks the tie. Thus by comparing the accuracy, precision, recall and F1 score, the best fit for our approach has to be Random Forest.

Chapter 6

Conclusion

It is effectively necessary to find a suitable teaching method for the children having autism spectrum disorder because every child with ASD has different characteristics and they should have specific education which is appropriate to their needs.

```
NAIVE BAYES

Confusion Matrix:
 [[28  0  6  0  0  0  8]
 [ 0 15  0  0  0  0  0]
 [ 0  0 24  0  0  0  0]
 [ 0  0  0 35  0  0  0]
 [ 0  0  0  0 18  0  0]
 [ 0  0  0  0  0  6  0]
 [ 0  0  0  0  0  0 13]]
Accuracy score:   90.84967320261438 %
Precision score:  91.70068027210885 %
Recall score:     95.23809523809523 %
F1 Score:         92.19421101774043 %
```

```
RANDOM FOREST

Confusion Matrix:
 [[42  0  0  0  0  0  0]
 [ 0 14  1  0  0  0  0]
 [ 0  0 24  0  0  0  0]
 [ 0  0  0 35  0  0  0]
 [ 0  0  0  0 18  0  0]
 [ 0  0  0  0  0  6  0]
 [ 1  0  0  0  0  0 12]]
Accuracy score:   98.69281045751634 %
Precision score:  99.09634551495017 %
Recall score:     97.94871794871796 %
F1 Score:         98.47634817473788 %
```

```
DECISION TREE

Confusion Matrix:
 [[42  0  0  0  0  0  0]
 [ 0 15  0  0  0  0  0]
 [ 0  0 24  0  0  0  0]
 [ 0  0  0 34  0  1  0]
 [ 0  0  0  1 17  0  0]
 [ 0  0  0  0  0  6  0]
 [ 0  0  0  0  0  0 13]]
Accuracy score:   98.69281045751634 %
Precision score:  97.55102040816325 %
Recall score:     98.79818594104309 %
F1 Score:         98.08477237048666 %
```

```
K NEAREST NEIGHBOURS

Confusion Matrix:
 [[41  0  0  0  0  0  1]
 [ 0 13  0  1  0  1  0]
 [ 0  0 22  0  0  0  2]
 [ 0  0  1 34  0  0  0]
 [ 0  0  0  1 17  0  0]
 [ 0  0  0  1  0  5  0]
 [ 1  0  0  0  0  0 12]]
Accuracy score:   94.11764705882352 %
Precision score:  92.64234953675945 %
Recall score:     91.88295831152973 %
F1 Score:         92.10401891252954 %
```

Figure 6.1: Algorithms and Scores

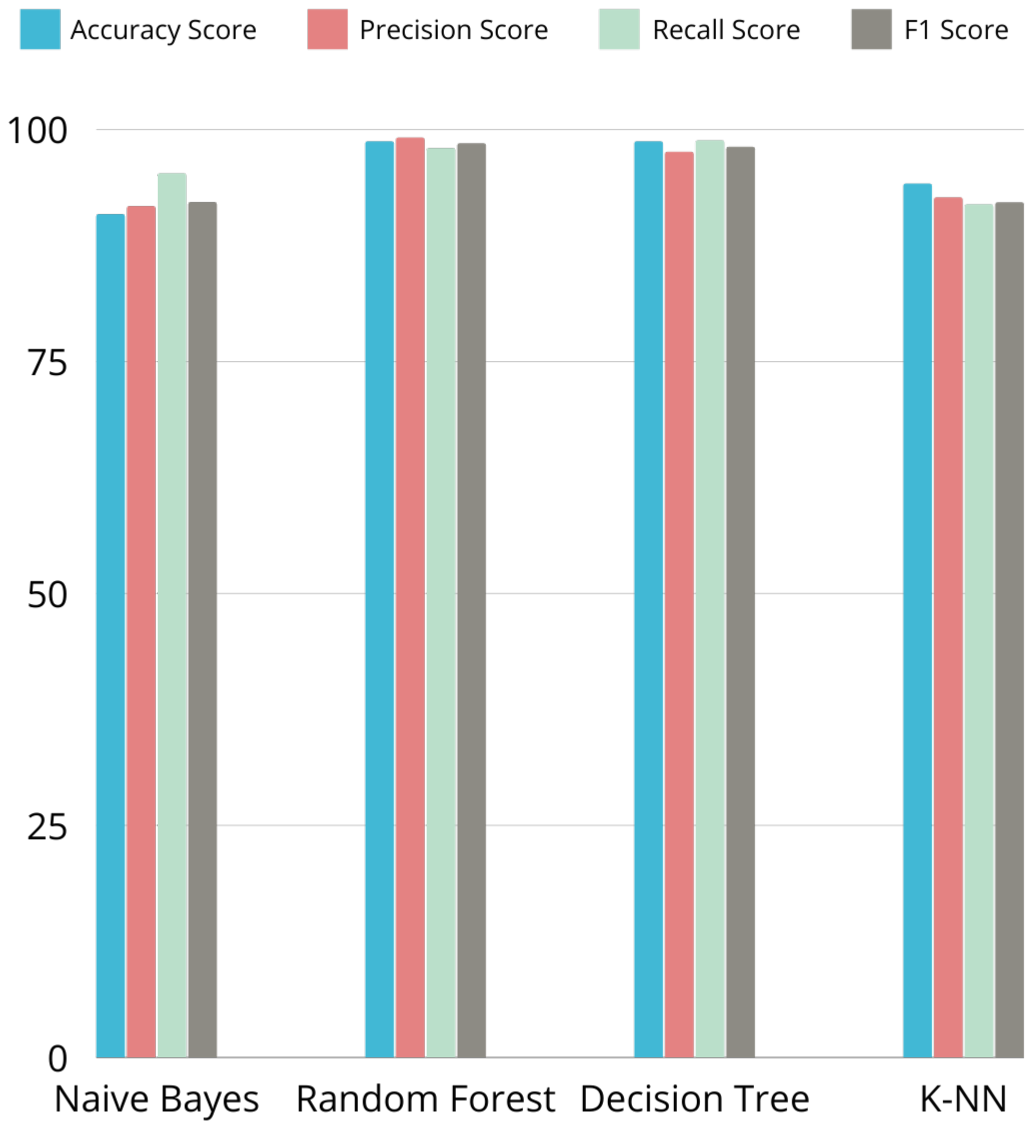


Figure 6.2: Bar diagram of Accuracy, Precision, Recall, F1 Score

Based on a report of World Health Organization (WHO), among every 160 children, at least one has been diagnosed with autism spectrum disorder (ASD). As per the article of Paulette Delgado, Autism Spectrum Disorder (ASD) in Education, Schools fail in order to fulfill the expectations of guardians by not recognizing or supporting the requirements of the children who need special treatment [42]. Social interactions are everywhere and changing continuously in schools. Moreover, some activity in the classroom may be suitable for general students but it can be unfitting for the children with ASD. One the other hand some social prompts or signs which indicate a child to change their certain behaviors are usually difficult to follow for an autistic child. There is confusion between teachers on how to treat an autistic child with special educational needs as there is a deficiency of the depth of knowledge. This might have a negative consequence on their education. Earlier studies have anticipated the challenges of teachers regarding the right education method. Not much research has been done previously to implement machine learning in identification of proper education needs of a child with ASD [45]. So, this research can be helpful for teachers, parents as well as education institutions to determine an effective learning method for an autistic child. We have suggested six education systems according to their necessity. After that four machine learning algorithms named Na¨ıve Bayes, Random Forest, Decision Tree and K-Nearest Neighbors have been used in this research. According to the result we have found that the Random Forest algorithm gave approximately 98.69% accuracy, 99.10% precision, 97.95% recall and 98.48% f1 score, which is the best fit in finding the most appropriate teaching approach based on the children's characteristics.

Bibliography

[1] U. G. Assembly, "Universal declaration of human rights," *UN General Assembly*, vol. 302, no. 2, pp. 14–25, 1948.

[2] P. M. Doyle, D. L. Gast, M. Wolery, M. J. Ault, and S. Meyer, "Teaching discrete skills to students with moderate mental retardation in small-group instructional arrangements," *Exceptionality: A Special Education Journal*, vol. 3, no. 4, pp. 233–253, 1992.

[3] P. Davies, "What is evidence-based education?" *British journal of educational studies*, vol. 47, no. 2, pp. 108–121, 1999.

[4] N. R. Council, *Educating children with autism.* National Academies Press, 2001.

[5] B. K. Hamre and R. C. Pianta, "Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade," *Child development*, vol. 72, no. 2, pp. 625–638, 2001.

[6] J. S. Handleman and S. L. Harris, *Preschool education programs for children with autism.* Citeseer, 2001.

[7] M. H. Charlop-Christy, M. Carpenter, L. Le, L. A. LeBlanc, and K. Kellet, "Using the picture exchange communication system (pecs) with children with autism: Assessment of pecs acquisition, speech, social-communicative behavior, and problem behavior," *Journal of applied behavior analysis*, vol. 35, no. 3, pp. 213–231, 2002.

[8] K. S. Langworthy-Lam, M. G. Aman, and M. E. Van Bourgondien, "Prevalence and patterns of use of psychoactive medicines in individuals with autism in the autism society of north carolina," *Journal of child and adolescent psychopharmacology*, vol. 12, no. 4, pp. 311–321, 2002.

[9] S. R. de Boer, D. E. Griswold, B. S. Myles, *et al.*, *Autism spectrum disorders: Interventions and treatments for children and youth.* Corwin press, 2004.

[10] C. I. Chesnevar, M. P. González, and A. G. Maguitman, "Didactic strategies for promoting significant learning in formal languages and automata theory," *ACM SIGCSE Bulletin*, vol. 36, no. 3, pp. 7–11, 2004.

[11] J. S. Handleman and S. L. Harris, *School-age education programs for children with autism.* Pro Ed, 2006.

[12] D. Hargreaves, "Teaching as a research-based profession: Possibilities and prospects (the teacher training agency lecture 1996)," *Educational research and evidence-based practice*, pp. 3–17, 2007.

[13] C. Tufvesson, *Concentration Difficulties in the School Environment-with focus on children with ADHD, Autism and Down's syndrome*. Lund University, 2007.

[14] S. Doumen, K. Verschueren, E. Buyse, V. Germeijs, K. Luyckx, and B. Soenens, "Reciprocal relations between teacher–child conflict and aggressive behavior in kindergarten: A three-wave longitudinal study," *Journal of Clinical Child & Adolescent Psychology*, vol. 37, no. 3, pp. 588–599, 2008.

[15] L. R. Holdheide and D. J. Reschly, "Teacher preparation to deliver inclusive services to students with disabilities: Tq connection issue paper.," *National Comprehensive Center for Teacher Quality*, 2008.

[16] J. Neitzel, "Overview of peer-mediated instruction and intervention for children and youth with autism spectrum disorders," *Chapel Hill, NC: National Professional Development Center on Autism Spectrum Disorders, Frank Porter Graham Child Development Institute, The University of North Carolina*, 2008.

[17] ——, "Steps for implementation: Antecedent-based interventions," *Chapel Hill, NC: The National Professional Development Center on Autism Spectrum Disorders, FPG Child Development Institute, The University of North Carolina*, 2009.

[18] E. Eldar, R. Talmor, and T. Wolf-Zukerman, "Successes and difficulties in the individual inclusion of children with autism spectrum disorder (asd) in the eyes of their coordinators," *International Journal of Inclusive Education*, vol. 14, no. 1, pp. 97–114, 2010.

[19] M. J. Morrier, K. L. Hess, and L. J. Heflin, "Teacher training for implementation of teaching strategies for students with autism spectrum disorders," *Teacher Education and Special Education*, vol. 34, no. 2, pp. 119–132, 2011.

[20] S. Aud, S. Wilkinson-Flicker, P. Kristapovich, A. Rathbun, X. Wang, and J. Zhang, "The condition of education 2013," *National Center for Education Statistics*, 2013.

[21] Y. Cai, N. K. Chia, D. Thalmann, N. K. Kee, J. Zheng, and N. M. Thalmann, "Design and development of a virtual dolphinarium for children with autism," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 21, no. 2, pp. 208–217, 2013.

[22] "Diagnostic and statistical manual of mental disorders," *Am Psychiatric Assoc (5th ed.)*, vol. 21, pp. 591–643, 2013.

[23] N. Dinuta, "Didactic strategies used in teaching–learninig of premathematical operations in preschool education," *Procedia-Social and Behavioral Sciences*, vol. 76, pp. 297–301, 2013.

[24] R. Verschuur, R. Didden, R. Lang, J. Sigafoos, and B. Huskens, "Pivotal response treatment for children with autism spectrum disorders: A systematic review," *Review Journal of Autism and Developmental Disorders*, vol. 1, no. 1, pp. 34–61, 2014.

[25] A. S. Eisenhower, J. Blacher, and H. H. Bush, "Longitudinal associations between externalizing problems and student–teacher relationship quality for young children with asd," *Research in Autism Spectrum Disorders*, vol. 9, pp. 163–173, 2015.

[26] A. S. Eisenhower, H. H. Bush, and J. Blacher, “Student-teacher relationships and early school adaptation of children with asd: A conceptual framework,” *Journal of Applied School Psychology*, vol. 31, no. 3, pp. 256–296, 2015.

[27] B. Caplan, M. Feldman, A. Eisenhower, and J. Blacher, “Student–teacher relationships for young children with autism spectrum disorder: Risk and protective factors,” *Journal of autism and developmental disorders*, vol. 46, no. 12, pp. 3653–3666, 2016.

[28] M. A. Alsehemi, M. M. Abousaadah, R. A. Sairafi, and M. M. Jan, “Public awareness of autism spectrum disorder,” *Neurosciences Journal*, vol. 22, no. 3, pp. 213–215, 2017.

[29] J. B. Leaf, J. H. Cihon, J. L. Ferguson, and S. M. Weinkauf, “An introduction to applied behavior analysis,” in Springer International Publishing, 2017, pp. 22–42.

[30] S. Hedges, “Technology-aided instruction & intervention (taii). ebp brief packet.,” *National Professional Development Center on Autism Spectrum Disorders*, 2018.

[31] F. F. Thabtah, *Autism screening data for toddlers*, Jul. 2018. [Online]. Available: https://www.kaggle.com/fabdelja/autism-screening-for-toddlers?select=Toddler+data+description.docx.

[32] L. E. Achenie, A. Scarpa, R. S. Factor, T. Wang, D. L. Robins, and D. S. McCrickard, “A machine learning strategy for autism screening in toddlers,” *Journal of developmental and behavioral pediatrics: JDBP*, vol. 40, no. 5, pp. 369–376, 2019.

[33] J. Blacher and B. L. Baker, “Collateral effects of youth disruptive behavior disorders on mothers’ psychological distress: Adolescents with autism spectrum disorder, intellectual disability, or typical development,” *Journal of autism and developmental disorders*, vol. 49, no. 7, pp. 2810–2821, 2019.

[34] B. Deepa and K. J. Marseline, “Exploration of autism spectrum disorder using classification algorithms,” *Procedia Computer Science*, vol. 165, pp. 143–150, 2019.

[35] N. Donges, “A complete guide to the random forest algorithm,” *Built In*, vol. 16, 2019.

[36] V. Kotu and B. Deshpande, “Chapter 5 - regression methods,” in *Data Science (Second Edition)*, V. Kotu and B. Deshpande, Eds., Second Edition, Morgan Kaufmann, 2019, pp. 165–197.

[37] I. C. Education, *Machine learning. ibm*, 2020.

[38] A. Network, *Community report on autism. cdc. gov*, 2020.

[39] C. Pratt and L. Steward, “Applied behavior analysis: The role of task analysis and chaining,” *Retrieved October*, vol. 15, p. 2020, 2020.

[40] S. Van Der Steen, C. H. Geveke, A. T. Steenbakkers, and H. W. Steenbeek, “Teaching students with autism spectrum disorders: What are the needs of educational professionals?” *Teaching and Teacher Education*, vol. 90, p. 103 036, 2020.

[41] T. Vos, S. S. Lim, C. Abbafati, *et al.*, "Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: A systematic analysis for the global burden of disease study 2019," *The Lancet*, vol. 396, no. 10258, pp. 1204–1222, 2020.

[42] P. Delgado, *Autism spectrum disorder (asd) in education*, Jul. 2021. [Online]. Available: https://observatory.tec.mx/edu-news/autism-spectrum-disorder-asd-in-education.

[43] F. Ertekin and B. Bağlama, "Examination of postgraduate thesis carried out with errorless teaching methods in individuals with autism spectrum disorder," *Near East University Online Journal of Education*, vol. 4, no. 1, pp. 109–117, 2021.

[44] V. Lampos, J. Mintz, and X. Qu, "An artificial intelligence approach for selecting effective teacher communication strategies in autism education," *npj Science of Learning*, vol. 6, no. 1, pp. 1–10, 2021.

[45] A. Cook and J. Ogden, "Challenges, strategies and self-efficacy of teachers supporting autistic pupils in contrasting school settings: A qualitative study," *European Journal of Special Needs Education*, vol. 37, no. 3, pp. 371–385, 2022.

[46] M. Fernandes, M. Homer, and J. Sano-Franchini, "Sneaky rhetorics, habitual media, and the affective politics of lurking on facebook,"

[47] GeeksforGeeks, *Decision tree*. [Online]. Available: https://www.geeksforgeeks.org/decision-tree/.

[48] JavaTPoint, *K-nearest neighbor (knn) algorithm for machine learning*. [Online]. Available: https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning.

[49] ——, *Machine learning tutorial: Machine learning with python*. [Online]. Available: https://www.javatpoint.com/machine-learning.

[50] S. Prabhakaran, *How naive bayes algorithm works? (with example and full code): Ml+*. [Online]. Available: https://www.machinelearningplus.com/predictive-modeling/how-naive-bayes-algorithm-works-with-example-and-full-code/.