International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Performance Analysis of the Impact of COVID-19 on Student Studies Using a K-NN Algorithm

Pratima Kumari *[1], Dileep Kumar Yadav [2], Sanjeev Gangwar[3]

Submitted: 15/01/2024 **Revised**: 23/02/2024 **Accepted**: 01/03/2024

Abstract: The education sector, being one of the largest sectors in the world is also affected by COVID-19 pandemic which result in an expected rise in dropouts, address such as a crisis and sustainability. To reduce this impact on the education system several questions arise that How to organize the education system and how the problem that has arisen can be eliminated? These questions were sent to the higher officials in the education system. Different types of suggestions are proposed, and they give justification to employ online learning to provide an individual with a relevant stress-free solution and to cope with the current situation.

Keywords: *Decision Tree, KNN, Gaussian Naïve Bayes, Case-based learning, ANN.*

1. Introduction

The impact of pandemic COVID-19 is observed in every sector around the world. It has affected educational systems worldwide, leading to near-total closures of schools, universities, and colleges. While countries are at different point in their COVID-19 infection rates, worldwide there are more than 1.2 billion children in the world have been severely impacted by temporary closure of educational institutions in order to deal with it educational institutions move towards blended mode of learning and resulted in increasing the digital literacy. Although it's still too early for much COVID-19-specific machine learning research to have been conducted and published, early experiments are promising. Furthermore, we can look at how machine learning is used in related areas and imagine how it could help with risk prediction for COVID-19.

Amid the COVID-19 epidemic, artificial intelligence will be taken into consideration as a tool for education. Artificial intelligence was created, added to machines, and seen as a development in technology. Using the methods offered by computer professionals, it enables the machine to carry out calculations and problem-solving. Artificial intelligence develops techniques and procedures that go beyond what the human mind is capable of. Artificial intelligence is still significant today because society still relies on machines to carry out tasks and procedures that make the system convenient and effective. Because it can carry out the vast majority of human duties in our society, this technology will therefore be increasingly pertinent during the current pandemic.

–[1]M.tech Scholar, Department of Computer Science and Engineering, VBSPU.
[1]Assistant Professor, Department of Computer Science and Engineering, VBSPU.
[2]Associate Professor, Department of Computer Science and Engineering VBSPU.
** Corresponding Author Email: pratimakumari160592@gmail.com*

- **Objective:** In this paper we create a model which can solve the problem in studies that arises due to COVID-19 and overcome that problem efficiently. So, in future if there is any problem that arises like this then we will be prepared to face that problem.
- To find out effectiveness and accuracy rate in K-NN, SVM, based clustering with students' studies measurement of online classes using machine learning. So, in future if there is any problem that arises like this then we will be prepared to face that problem. Classification and clustering with study in disaster, we can use data mining and machine learning algorithm for optimal selection better compare accuracy to various machine learning algorithms.
  Motivation: In this paper there are several problems that arise due to COVID-19 especially in the education sector and the study of students is severally disturbed by this problem. That's why we wanted to create a model which helps students to overcome the problems that arise during their studies.
  Contribution: In this paper we performance d several machine leaning algorithms to obtain better accuracy than previous research. We work on a dataset which contains several cases, and we study them through machine learning algorithms.

2. Literature Survey

Literature review of the research in according to the scientist in research paper at COVID-19 Impact on student's education Using Machine Learning.

Monica Ciolacu a et al. [1] In this work, we introduce a novel strategy for fostering AI in Education 4.0. The AI-assisted higher education process with wearable technology and smart sensors is our first contribution to the field. We next go over our initial findings from the implementation of

didactic techniques for Education 4.0 using learning analytics and machine learning algorithms.

Rachmawati Findiana et al. [2] The results have not been optimized because the selection of the students is still done manually. Students with strong academic records may not be admitted, and vice versa; students with ordinary academic records may still be admitted to General Higher Education using this method. Consequently, in order to get the best outcomes, a classification system is required to forecast student graduation at Public Higher Education via this path. The classification of students graduating from SNMPTN, SPAN PTKIN, SMPN, and not received / not passed will be made in this study using the parameters of the average semester report card grades 1, average semester report card grades 2, average semester report card grades 3, grades average semester report card 4, and the average grade report card semester 5.Because the kids still have to be manually included in the selection process, the outcomes have not been improved. Students with strong academic records may not be admitted, and vice versa; however, this pathway may also be used by students with ordinary academic records. To achieve the best results, a classification system is necessary to forecast student graduation at public higher education using this approach. In this study, the classification of student graduation at SNMPTN, SPAN PTKIN, SNMPN, and not received / not passed will be made using the parameters of the average semester report cards grades 1, 2, 3, and 4, as well as the average grade report cards semester 5.

Marinela mircea et al. [3] Information and communications technology (ICT) advancements bring about change in many spheres of life, including the educational system.

Safika Mohd Shafiee et al. [4] Students in higher education are among the groups who are most at risk of developing mental health issues. To assist a person with a mental health problem, it is challenging and difficult to pinpoint the causes of those problems. This study aims to: (1) assess the mental health issues that higher education students face; (2) identify the contributing elements; and (3) review the machine learning tools that are now available for analyzing and forecasting these issues. The findings of the work will be applied to other research to advance the discussion on mental health issues for use in computer modeling. and depending on the measure come, a rival is the support vector machine, which is the mean squared error.

Indy Ho et al. [12] There is a lack of knowledge regarding the determinants predicting student satisfaction for this unique learning environment in crisis in higher education during the COVID-19 epidemic. The primary learning tools used in this study were Microsoft Team and Moodle, and the participants were undergraduate students (N = 425) from a variety of disciplines at a self-funded institution in Hong Kong. In confirming the occurrence of varied student assessment ratings for teachers in gender and socioeconomically diverse higher educational environments, this paper adds to the body of literature. This study investigated whether students' perceptions of their teachers in higher education across a range of subjects are influenced by disparities in gender and socioeconomic status among both teachers and students.

Rodrigo Pessoa Medeiros et al. [6] conducted a study to assess the effectiveness of deep learning methods in identifying false information pertaining to the COVID-19 pandemic. Their study attempted to fulfil the pressing need for precise identification of disinformation within this worldwide health problem. Wani et al. conducted an analysis of several deep learning models to get insights into how well they function and how they might be used to stop the spread of false news during the COVID-19 epidemic. By tackling beginning programming in a more modern setting, the paper expands on the findings of earlier systematic literature surveys. It suggests categorizing the difficulties in learning and teaching beginning programming in higher education and emphasizes important areas for a research roadmap.

Gaurav Bhatia et al. [7] The design of a private cloud for higher education is presented in this paper, along with a technique for implementing a proof-of-concept using the Open Stack platform. In order to make building the private cloud easier, a model for measuring the number of physical resources needed for a given number of virtual machines has also been developed.

Chitra Jalota et al. [8] This study will examine numerous data mining strategies that are helpful for estimating students' performance levels. To do this, we applied the Callboard 360 dataset to Weka to analyze data mining approaches.

Mircea E. Ardeleanu et al. [9] It has been discovered that many young individuals quit out of higher education after a year or two of studies or eventually feel disappointed in their careers due to an improper oriented choice of vocation. Counseling and guiding high school graduates toward postsecondary education and pointing them in the direction of a professional career that is appropriate for their skills and abilities is a complex task that, sadly, does not seem to receive the attention and importance it deserves.

José Luis Martín Núñez et al. [10] This paper examines the process of converting a traditional course into an open course over four academic years, revealing improvements in student satisfaction and assessment views. It also highlights risks associated with massive online courses.

Gabriela Cejas et al. [11] Baseline 2020 investigates ongoing research projects within the Faculty of Engineering, examining diversity, multidisciplinary, and social relevance. The study reveals an imbalance between

project funds and investments, and a lack of involvement in funding initiatives. The organization prioritizes alternative funding sources and encourages student participation.

MiaoBa et al. [12] The Changshu Institute of Technology and the University of Wisconsin-Eau Claire have collaborated since 2018 to establish a teaching mode of assessment at ordinary times (AOT) based on the outcome-based education (OBE) concept, aiming to improve the EEA system and ensure success in the "Material Science and Engineering" course.

3. Methodology

3.1 Machine learning models

•**Dataset:** Machine learning datasets come in a wide variety of forms. Text, audio, video, and image data are a few of the most popular types. The use cases for each type of data are different from one another. For systems that need to comprehend natural language, text data is a fantastic option.

•**Training dataset:** To train a machine learning model, one uses training data, a large dataset. Using training data, prediction models that use machine learning algorithms are taught how to extract characteristics that are relevant to certain business goals. For supervised ML models, the training data is labeled.

•**Test data-set:** Using a subset of the training dataset known as the test dataset, a final model is evaluated objectively. Depending on the validation dataset, there are additional methods for calculating an unbiased or biased assessment of a model's ability to handle unknown input.

•**Testing model:** A fully trained model's performance on a testing set is evaluated through the model testing procedure in machine learning.

•**Generation of model:** A generative model is a kind of machine learning model that seeks to produce new, comparable data by understanding the underlying patterns or distributions of the original data.

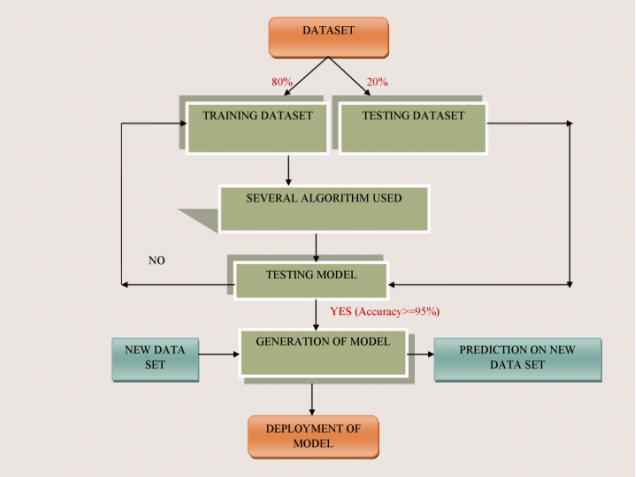

Fig.1. Working Diagram of Machine Learning Models

In machine learning, generative modeling is a sort of unsup ervised learning in which regularities or patterns in the inp ut data are automatically found and learned so that the mod el can be used to generate or output new examples that cou ld have been realistically taken from the original dataset. Generative models are different from discriminative models. Discriminative models discriminate between different kinds of data instances, while generative models can generate new data instances. For example, a generative model could generate new photos of animals that look like real animals, while a discriminative model could tell a dog from a cat.

•**Model deployment :**Model deployment in machine learning is the process of integrating a trained model into an existing production environment where it can take in an input and return an output. The goal is to make the predication from our trained machine learning model available to others.

There are several ways to deploy a machine learning model, including cloud-based platforms such as Algorithmic and Python Anywhere. These platforms provide a simple and faster way to deploy our machine learning model into production.

The four steps to machine learning deployment include:

•Develop and create a model in a training environment.

•Test and clean the code ready for deployment.

•Prepare for container deployment.

•**Prediction new data:** In machine learning, predicting new data is the process of using a trained model to make predictions on new data instances. Once we have trained a machine learning model, we can use it to make predictions on new data instances that the model has not seen before. To make predictions on new data, we can use the predict () function in scikit-learn, which is a popular machine learning library in python. The predict () function takes in new data instance and returns the predicted class or value for each instance.

Here are the steps to predict new data using scikit-learn.

•Train a machine learning model on a training dataset.

•Prepare the new data instances that you want to make predictions on.

•Call the predict () function on the trained model with the new data instances as input.

•Use the predicted values to make decisions or take actions.

3.2 Machine learning based algorithm

•**Data cleaning:** This involves removing or correcting any missing incomplete or incorrect data.

•**Data transformation:** In order to do this, the data must be transformed into a more practical format, for as by turning categorical data into numerical data.

•**Feature selection:** This entails deciding on the feature or viable that will be used in the analysis as being the most pertinent.

•**Feature engineering**: In order to increase the analysis's accuracy, additional features or variables are created from the data that is already available.

•**Data integration**: For the purpose of creating a single data set that can be evaluated collectively, this entails merging data from many sources.

•**Data reduction:** As a result, the information must be compressed using methods like principal component analysis or reduced in size by eliminating occurrences or features that are unnecessary.

Overall preprocessing is a crucial step in the data analysis and machine learning process, as it can significantly impact the accuracy and effectiveness of the analysis.

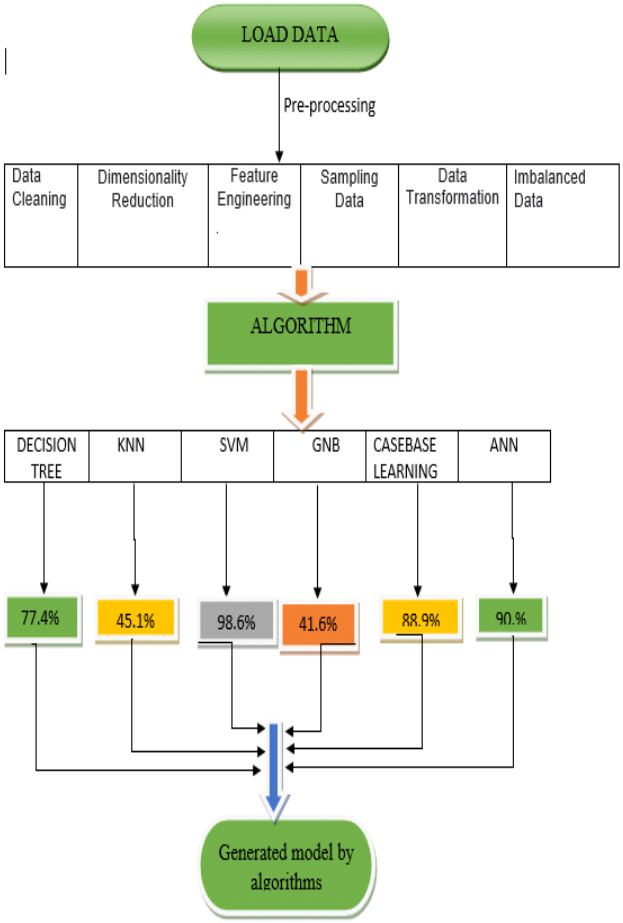

Fig.2. Block Diagram of Model Generation

3.3 Algorithm

A finite number of steps that must be taken in order to solve a mathematical problem, or a set of finite rules or instructions to be complied with in calculations or other problem-solving operations. machine learning algorithms are computational models that Without the need for explicit programming, let computers recognize patterns, forecast, or form opinions based on data. These algorithms serve as the cornerstone of contemporary artificial intelligence and find use in a multitude of contexts, such as recommendation engines, fraud detection, image and speech recognition, natural language processing, and autonomous vehicles.

3.4 K-NN (K-Nearest Neighbor)

The k-nearest neighbor algorithm is the most widely used supervised learning technique; although it is typically used solely for categorization tasks, it may also be used for regression. Each new training sample is compared to previous training, data for classification in the K-NN algorithm. KNN predicts the values of new data points using "feature similarity"; each new data point is given a value based on how similar it is to prior data points that have been stored in memory.

•The K-NN method places the new case in the category most comparable to the existing categories based on the assumption that the new case and data are similar to the examples that are already available.

•The K-NN algorithm classifies a new data point based on similarity after storing all the relevant data. This indicates that the K-NN algorithm can quickly classify newly appearing data into a well-suited category.

•Although the K-NN technique is mostly utilized for classification problems, it can also be employed for regression.

•K-NN does not make any assumptions about the underlying data because it is a non-parametric algorithm.

The algorithm is also known as a lazy learner because, rather than learning straight away from the training set, it saves the dataset and uses it to execute an action when classifying data. During the training phase, the KNN algorithm simply saves the dataset and categorizes newly received data into groups that are strikingly similar to each other.

Algorithm 1: K-NN (K-Nearest Neighbor)		
Input	:	dataset
Step 1	:	Load the data set
Step 2	:	Choose only value of nearest neighbour data points (k). k can be any integer value (say k=5)
Step 3	:	Calculate the Euclidean distance between test data (query instance, x) and each row of training data. Note down the value of Euclidean distance in an ordered table.
Step 4	:	Arrange the Euclidean distance table in ascending order. Choose the top K rows (say k=5) forum this table.
Step 5	:	Now assign a class to the new instance point query (test data) based on nearest class of rows. Thus,

		assign the new data points to that category (class) for which the member of neighbor is maximum.
Step 6	**:**	End
Euclidean distance ((X2-X1)2+ (Y2-Y1)2)1/2		

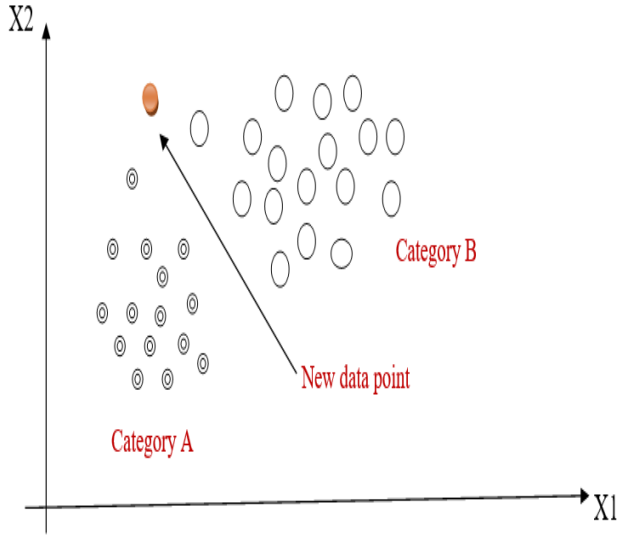

Fig.3. K-Nearest Neighbor

3.4 Support Vector Machine

SVM is a very well-liked supervised method that is utilized for both regression and classification tasks. The goal of an SVM algorithm is to find a hyper plane in an N-dimensional spaced that clearly classifies the input points; nevertheless, SVM is primarily employed for classification.

The Support Vector Machine (SVM) learning algorithm is presented in this set of notes. SVMs are some of the best "off-the-shelf" supervised learning algorithms—many people even argue that they are the best. In order to explain the SVM story, we must first discuss margins and the notion of separating data that has a significant "gap." We will then stray into the topic of Lagrange duality as we discuss the optimal margin classifier.

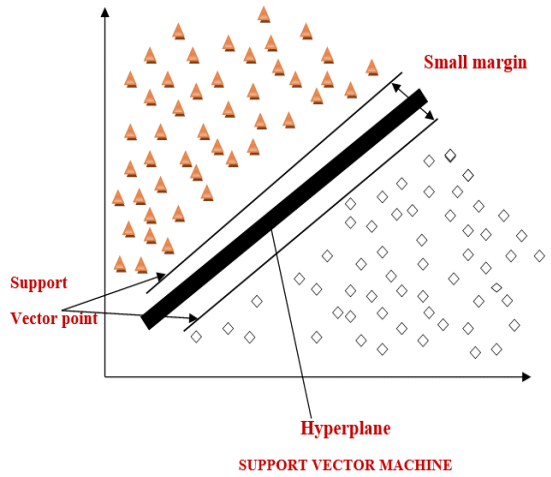

Fig.4. Support Vector Machine

Support vectors are the nearest data points to the hyper plane that have an impact on the hyper plane's position. In the plane, every data point is referred to as a vector (point).

Margin is the distance between two lines drawn from the nearest data points of various closes.

3.5 ANN (Artificial Neural Network)

ANN was inspired by the human brain's information-processing model. The human brain is made up of billions of interconnected neurons. Each neuron takes in information from other neurons, processes it, becomes stimulated, and then passes it along to other neurons.

Fig.5. Basic ANN Model

3.6 Gaussian Naïve Bayes (GNB)

Gaussian naïve Bayes (GNB) is a classification algorithm that is based on Bayes' theorem. It is a type of naïve Bayes algorithm that assumes that the features of a dataset are independent of each other, and that they have a Gaussian distribution. In GNB, the probability of a sample elonging to a particular class is calculated using Bayes' theorem which states that:

$P(y/x) = p(x/y)*p(y)/p(x)$

Where

P(y/x) is the probability of class y given the features x

P(x/y) is the probability of the features x given the class y

P(y) is the prior probability of class y

P(x) is the prior probability of the features x

GNB assumes that the features x are independent of each other given the class y, which means that:

$P(y/x) = p(x1/y)*p(x2)/p(y)* \ldots * p(xn/y)*$

Where

X1, X2, …xn are the features of the sample

p(xi/y) is the probability distribution of the feature xi given the class y

GNB assumes that the probability distribution of each feature xi given the class y is Gaussian (normal), which means that:

P (xi/y) = (1/sqrt (2*pi*sigma^sigma^2))*exp (-((xi-mean) ^2)/ (2*sigma^2))

Where:

Mean is the mean of the feature xi for class y

Sigma^2 is the variance of the features xi for class y

By using these assumptions, GNB calculates the probability of a sample belonging to a particular class, and then assigns the sample to the class with the highest probability.

GNB is often used in text classification and spam filtering, as well as in other applications where the features can be modeled as continuous variables with a Gaussian distribution.

The naïve bays rule can apply for a method for going from P(X/Y), it is known by the preparation of the dataset to discover P(X/Y).What occurs if Y has multiple classes? We process the likelihood of each class of Y and let the most elevated success.

P(X/Y) = P (X∩Y)/P(Y)[P (Evidence / outcome) (Known from training data)]

P(Y/X) = P (X∩Y)/P(X)[P (Outcome / Evidence) (to be predicted for test data)]

Naive Bays calculations are an arranging approach based on Bays' hypothesis and the assumption that each of the indicators is independent of the others. The governing aim in naive Bayesian depiction is to identify the back probabilities, such as the probability of a name given some watched feature, (L | features).

3.7 Naive Bayes Classifiers

A group of classification algorithms based on Bayes' Theorem are known as naive Bayes classifiers. It is actually a family of algorithms rather than a single algorithm, and they are all based on the same principle—that is, each pair of features being classified stands alone. Let's start by thinking about a dataset. The Naïve Bayes classifier, one of the most straightforward and efficient classification algorithms, facilitates the quick creation of machine learning models with quick prediction capabilities.

•Easy to implement and computationally efficient.

•Effective in cases with a large number of features.

•Performs well even with limited training data.

The advantages of the optimized J48 classifier include:

•Spam Email Filtering: Classifies emails as spam or non-spam based on features.

•Text Classification: Used in sentiment analysis, document categorization, and topic classification.

•Medical Diagnosis: Helps in predicting the likelihood of a disease based on symptoms.

•Credit Scoring: Evaluates creditworthiness of individuals for loan approval.

•Weather Prediction: Classifies weather conditions based on various factors.

3.8 Case –Based Learning

Case-based reasoning (CBR) is used for classification and regression. Case-based reasoning is the process of figuring out how to solve brand-new issues. The CBR is a sophisticated instance-based learning technique used to resolve more challenging issues. The Euclidean distance metric is not used. A similar case is first checked in memory when a new case needs to be classified. The solution to the case is also retrieved if any similar cases are discovered in the memory cache. Draw information from memory. Look for any previous solutions to the issues at hand. Using the knowledge gained, offer a solution. Change it to suit the needs of the new circumstance. Consider the solution's application in a new setting. Save this novel approach to problem-solving to your memory system.

With case-based learning, students deal with scenarios that mimic particular real-life occurrences to learn how to interact with and modify fundamental knowledge. A single-sentence physics word problem, a nursing case the size of a textbook, or a case that takes up an entire semester in a law school are instances of case-based learning.

Cases are typically narratives or stories with characters and situations. These stories are told through interactive cases, videos, written texts, and even games and simulations. Positive attributes of instances include their story-telling ability, versatility, and effectiveness in fostering independent learning.

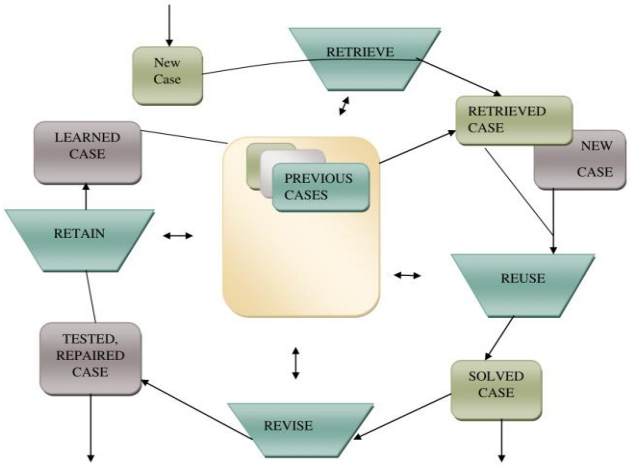

Fig.6. Structure of Case-based Learning Process

3.9 Decision Tree Algorithm

Classification and regression problems can be resolved using supervised learning techniques called decision trees, albeit they are usually best suited for them. This is a tree-structured classifier, where each leaf node represents the classification outcome, internal nodes represent the features

of a dataset, and branching represents the decision-making process. The test is mainly utilized in classification techniques and is conducted utilizing the characteristic of the dataset that is provided.

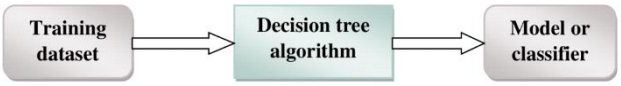

Fig.7. Block diagram of Decision support model

•Although decision trees are a supervised learning technique, they are mostly employed to solve classification problems. However, they can also be used to solve regression problems. This classifier is tree-structured, with internal nodes standing in for dataset attributes, branches for decision rules, and leaf nodes for each outcome.

•The Decision Node and the Leaf Node are the two nodes that make up a decision tree. While leaf nodes represent the result of decisions and do not have any more branches, decision nodes are used to make any kind of decision and have numerous branches.

•The decisions or the test are performed based on features of the given dataset.

•It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions.

•It is called a decision tree because, similar to a tree, it starts with the root node, which expands on further branches and constructs a tree-like structure.

•In order to build a tree, we use the CART algorithm, which stands for Classification and Regression Tree algorithm.

•A decision tree simply asks a question and based on the answer (Yes/No), it further splits the tree into sub-trees.

The procedure in a decision tree begins at the root node in order to forecast the class of the given dataset. This algorithm follows the branch and advances to the next node by comparing the values of the root attribute with the record (actual dataset) attribute.

The method proceeds to the next node by comparing its attribute value with those of the other sub-nodes once more. It keeps doing this till it gets to the tree's leaf node. The following algorithm can help you better understand the entire process.

- **Improved performance:** By optimizing the classifier's parameters, it is possible to achieve higher accuracy and better overall performance on the given dataset.
- **Flexibility:** The optimization process allows for customization of the classifier to better suit the characteristics of the dataset and the problem at hand.
- **Robustness:** By considering different options and performing cross-validation, the optimized Naive Bayes classifier can potentially handle variations and uncertainties in the data more effectively.
- **Generalizability:** The optimized classifier is trained to perform well on the training dataset and is expected to generalize well to unseen data, making it a reliable predictive model.

Python Implementation of Decision Tree

Now we will implement the Decision Tree using Python. For this, we will use the dataset "covid.csv," which we have used in previous classification models. By using the same dataset, we can compare the Decision tree classifier with other classification models such as KNN SVM, decision tree, G.N.B., ANN, Case-Base Learning, and Decision Tree method etc.

•Steps will also remain the same, which are given below.
•Data Pre-processing step
•Fitting a Decision-Tree algorithm to the Training set.
•Predicting the test result.
•Test accuracy of the result (Creation of Confusion matrix)
•Visualizing the test set result.

4. Implementation

4.1 Dataset

covid.csv: The entire dataset regarding student details with study media has been gathered from the Kaggle, WEKA and UCI repositories in 1131entities with 23columns.

4.2 Number of Attributes

23 (Id, Region of residence, age of subject, time spent online, Rating of online class experience, 23 time utilize).

The file is named "submit.csv". An example submission that you may use.

Dataset link: https://www.kaggle.com/

5. Experimental Results and Discussions

Understanding the influence of test data on online classes outcomes, as well as the connection between qualities, is the focus of this research study on Machine Learning Algorithms on performance of investigation on student study Data in covid-19. Set the accuracy and training score to 98.6 percent. Who has unending pandemic and has not enervated the algorithm on the basis of the computation situated among the KNN, SVM, G.N.B., ANN, Case-Base Learning, and Decision Tree method utilized in this proceeding? Support vector machine has the greatest accuracy of 96.61 percent, while Decision Tree has 77.4 percent accuracy. Case-Base Learning has a high Training Score of 88.9 percent, while Decision Tree has a high Training Score of 77.4 percent.

In future there may be another pandemic happen in future or anything else that can cause or affect the studies of students like this (covid-19) so in future we will be ready to face problems like this and overcome with this problem. And in

future we can make this model better using AI and several other techniques to enhance this model. The accuracy and training score of the machine learning algorithm for forecasting students' study habits. The study will employ additional machine learning algorithms, such as ANN and K-Means cluster algorithms, regression analysis and cluster analysis algorithms, to improve accuracy and training score.

Four assessment metrics—accuracy, precision, recall, and support—are used to track and evaluate the K-NN algorithm's performance. A key metric that gauges the proportion of precise forecasts the algorithm produces is accuracy. The assessment takes into account cases that are successfully recognized as true as well as instances that are correctly identified as false, or true positives and true negatives, and compares them to the total number of predictions produced. Better performance is correlated with higher accuracy levels. It is described as:

Accuracy = (true positives + true negatives) / (true positives + true negatives + false positives + false negatives) (1)

The number of true positives divided by the total number of positive predictions yields the precision. A reduced number of false positives is correlated with a better accuracy level. Formally, it is defined as:

Precision = true positives / (true positives + false positives) (2)

The number of true positives divided by the total number of genuine positives in the dataset yields the recall. There are fewer false negatives when the recall value is larger. Formally, it is defined as:

Recall = true positives / (true positives + false negatives) (3)

Table 1: Performance Comparison of Classifiers using Accuracy, Precision, Recall, and support

Algorithm	**Averages**	**Precision**	**Recall**	**Support**
K-NN	Macro Averages	0.74	0.65	114
ANN	Macro Averages	0.29	0.50	114
	Weighted Averages	0.35	0.59	
Decision Tree	Macro Averages	0.94	0.95	114
Naive Bayes	Macro Averages	0.63	0.51	114
	Weighted Averages	0.62	0.60	114
Case-based learning	Macro Averages	0.96	0.96	114
	Weighted Averages	0.96	0.96	114
Support Vector Machine	Macro Averages	0.29	0.50	114
	Weighted Averages	0.35	0.59	114

Results of Prediction, Support while using various algorithms with Delhi NCR students

Table 1 lists performance metrics for a variety of deep learning and machine learning algorithms, such as Naive Bayes, K-Nearest Neighbor, Case-based learning, Support Vector Machine, ANN, and Decision Tree, that are assessed using four metrics in order to identify potentially false information in particular contexts.

1. The model's accuracy is measured by comparing the number of accurately predicted instances to the total instances, indicating its proficiency in accurately categorizing.
2. Precision is a metric that measures the proportion of accurately predicted positive observations out of the total number of expected positives, indicating a high level of accuracy and a low occurrence of false positives.
3. Recall, or sensitivity, is the proportion of predicted positive observations to actual positive observations, indicating the comprehensiveness of a classifier. A high recall value indicates successful retrieval of relevant

results.

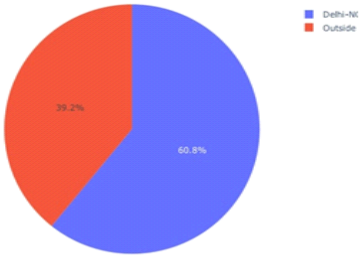

Fig.8. Result Maximum no. of students from which region most

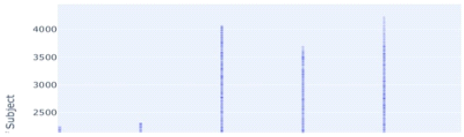

Fig.9. Result Time Spent on Different subject on online Class

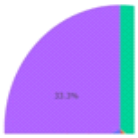

Fig.10. Result Rating of online Classes

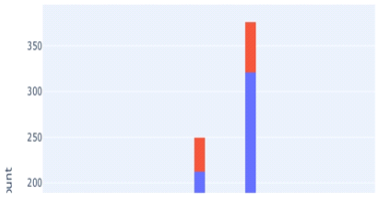

Fig.11. Result Time spent on self-study

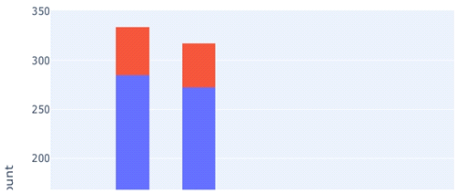

Fig.12. Result Time spent on Social Media

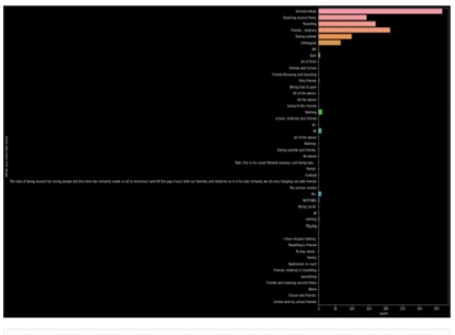

Fig.13. Result Stress busters

6. Conclusion

It is beyond dispute that the harmful grip COVID-19 pandemic has completely destroyed the global educational system. As a result, in predicting the future resonance of educational institutions, which are inextricably linked to the lives of students and teachers, policymakers should take the role of artificial intelligence (AI). ML algorithms can play a significant part in settling the nation's pillar. An effective predictive model is built and compared with various features in this investigation of the influence of COVID-19. Support Vector Machine, and Case-based learning provided the best accuracy among all, at 98.6%. AI can therefore help develop better policies and provide strong incentives to support the welfare of humans.

A system called Information Cultivate is used to investigate the region. Naive Bayes, K-Nearest Neighbor, Case-based learning, Support Vector Machine, ANN, and Decision Tree were among the machine learning techniques used in the study. Decision Tree and Artificial Neural Network techniques are employed to forecast which students learn from COVID-19 disasters and which do not. The most current data indicates that the Case-based learning classifier outperforms other models in terms of accuracy and execution time when it comes to predicting the best show outcome.

In order to carry out analysis several ML model used which give following Accuracy:

S.No	Algorithm Used	Training Accuracy Obtained	Testing Accuracy Obtained
1	K-Nearest Neighbor	50.6%	45.1%
2	Decision Tree	82.6%	77.4%
3	GaussianNaïve Bayes	47.8%	41.6%
4	SupportVector Machine	99.8%	98.6%

5	Case-Based Learning	93.1%	88.9%
6	ArtificialNeural Network	86.2%	81.2%

Table 2 – Find out of the Training Score and Accuracy help of the Machine Learning Algorithms

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Monica Ciolacu, Ali Fallah Tehrani, Leon Binder, and Paul Mugur Svasta, "Education 4.0 – Artificial Intelligence assisted Higher Education: Early recognition System with Machine Learning to support Students", Success 25-28OCT 2018.

[2] RachmawatiFindiana Eko MulyantoYuniarno Endroyono, "Classification of Graduates tudent on Entrance Selection Public Higher Education Through Report Card Grade Path Using Support Vector Machine Method", 2020IEEE .

[3] Marinela mircea, marian stoica,and bogdan ghilic-micu," Investigating the Impact of the Internet of Thingsin Higher Education Environment", 29 jan 2021

[4] Nor Safika Mohd Shafiee, SofianitaMutalib, "Prediction of Mental Health Problems among Higher Education Student Using Machine Learning", 1 june2020.

[5] Indy Ho, Kai-Yuen Cheong, Anthony Weldon, "Predicting student satisfaction of emergency remote learning in higher education during COVID-19 using machine learning techniques", april 2021.

[6] Rodrigo Pessoa Medeiros, Geber Lisboa Ramalho, and Taciana Pontual Falcão, "A Systematic Literature Review on Teaching and Learning Introductory Programming in Higher Education", 2018 IEEE.

[7] Gaurav Bhatia, Ismail Al Noutaki, Sheikha Al Ruzeiqi, Jawaher Al Maskari, "Design and Implementation of Private Cloud for Higher Education using OpenStack", 2018IEEE.

[8] Chitra Jalota, Rashmi Agrawal, "Analysis of Educational Data Mining using Classification", IEEE 2019.

[9] Anika Gupta, Deepak Garg, Senior Member, IEEE, and Parteek Kumar, "Analysis of Students' Ratings of Teaching Quality to Understand the Role of Gender and Socio-Economic Diversity in Higher Education", 2020IEEE.

[10] Mircea E. Ardeleanu, Dan G. St_nescu, " A New Research Concerning Some InfluenceFactors in the Orientation of Highschool Graduates Towards Higher Technical Education", IEEE 2016.

[11] José Luis Martín Núñez, Member, IEEE, Edmundo Tovar Caro, Senior Member, IEEE,and José Ramón Hilera González, "From Higher Education to Open Education: Challenges in the Transformation of an Online Traditional Course", IEEE 2016.

[12] DSc. Gabriela Cejas, MSc. Juan José Vulcano, MSc. Ariel Uema, Eng. Fernando Camuzzi "Baseline 2020 in engineering research projects in a private Institution of Higher Education", IEEE 2020.

[13] MiaoBa, Jun-jun Kong, Yu-feng Wang, Long-fei Liu, "A teaching mode in Sino-US cooperation in higher Education", IEEE2021.

[14] Abir Abdullha, Sheikh Abujar, "COVID-19 Data Analysis and the situation predication using machine learning based on Bangladesh prespective", 2020.

[15] safaBahri,Moetezkdayem, Nesrine Zoghlami, "Deep Learning for COVID-19 prediction", 2020.

[16] Akshay Kumar Siddhu, Dr. Ashok Kumar, Dr. Shakti Kundu, "Review Paper for Detection of COVID-19 from Medical Images and/ or Symptoms of patient using Machine Learning Approaches", 2020.

[17] "Machine Learning Algorithms for Forecasting COVID-19 Confirmed Cases in America", 2020.

[18] Zainab Abbas Abdulhussein Alwaeli, Asst. Prof. Dr. Abdullahi Abdu Ibrahim "Predicting COVID-19 Trajectory Using Machine Learning",2020'.

[19] R. Martínez, O. Alvarez-Xochihua, O. Mejia, A. Jordan, and J. González-Fraga "Use of Machine Learning to Measure the Influence of Behavioral and Personality Factors on Academic Performance of Higher Education Students" IEEE LATIN AMERICA TRANSACTIONS, VOL. 17, NO. 4, APRIL 2019".

[20] Mircea E. Ardeleanu, Dan G. Stanescu, "A New Research Concerning Some Influence Factors in the Orientation of Highschool Graduates Towards Higher Technical Education"

[21] Ulrike Baumol, Alina Bockshecker " Evolutionary change of higher education driven by digitalization"

[22] DSc. Gabriela Cejas, MSc. Ariel Uema, MSc. Juan Jose Vulcano, Eng. Fernando Camuzzi, "Baseline 2020 in engineering research projects in a private Institution of Higher Education", 2020.

[23] Fahim Faisal; Mirza Muntasir Nishat; Md. Ashif Mahbub; Md. Minhajul Islam Shawon; Md. Mahbub-Ul-Huq Alvi, "Covid-19 and its impact on school

closures: a predictive analysis using machine learning algorithms",2021.

[24] Iman Akour , PhD Muhammad Alshurideh, PhD Barween Al Kurdi , PhD AmelAlAli, "Using Machine Learning Algorithms to Predict People's Intention to Use Mobile Learning Platforms During the COVID-19 Pandemic: Machine Learning Approach",2022.

[25] Sheikh Mufrad Hossain, Md. Mahfujur Rahman, Alistair Barros & Md.Whaiduzzaman, "Impact Prediction of Online Education During COVID-19 Using Machine Learning: A Case Study".

[26] Asmat Ara Shaikh, Anuj Kumar, Kruti Jani, Saloni Mitra, Diego A. García-Tadeo, Agilandeswari Devarajan, "The Role of Machine Learning and Artificial Intelligence for making a Digital Classroom and its sustainable Impact on Education during Covid-19",2023.

[27] José Jesús Sánchez Amate, Antonio Luque de la Rosa, Rafaela Gutiérrez Cáceres and Alejandro Vargas Serrano, "The Effects of COVID-19 in the Learning Process of Primary School Students: A Systematic Review",2022.