International Journal of Multidisciplinary Engineering in Current Research - IJMEC
Volume 9, Issue 8, August-2024, http://ijmec.com/, ISSN: 2456-4265

DETECTION OF MENTAL STATE FROM EEG SIGNAL DATA: AN INVESTIGATION WITH MACHINE LEARNING CLASSIFIERS

Mohammed Haroon Hussain[1], Syed Shah Maaz Uddin[2], Syed Shah Quzaima Hassan[3], M.Neelima [4]
[1,2,3] B.E. Student, Department of IT, Lords Institute of Engineering and Technology, Hyderabad
[4] Assistant Professor, Department of IT, Lords Institute of Engineering and Technology, Hyderabad
mneelima@lords.ac.in

Abstract

The mental state of an individual is influenced by a complex interplay of neural activities, shaped by numerous external and internal factors. By analyzing EEG patterns, it is possible to ascertain an individual's mental state. Utilizing a dataset from Kaggle, ten machine learning techniques were explored, and models were developed. Hyperparameter tuning was performed using the RandomSearchCV method, and a comparative analysis was conducted for both tuned and non-tuned hyperparameters. Upon evaluating the performance metrics, the Support Vector Machine (SVM) demonstrated the highest accuracy at 95.36%. Gradient Boosting (GrB) also showed a promising accuracy of 95.24%, while K-Nearest Neighbors (KNN) and XGBoost (XGB) both achieved an accuracy of 93.10%. Consequently, the effective integration of ML-based detection methods can potentially regulate a person's mental state, fostering a deeper understanding of human psychology and predicting their behavior.

I. INTRODUCTION

Mental state refers to the condition of one's mind, which can be understood from various perspectives such as consciousness-based, intentionality-based, and functionalism, as indicated by numerous studies and research [1]. Automated detection of people's mental states can serve as an effective tool to monitor inappropriate activities (e.g., drowsy drivers), mental health issues (e.g., anxiety, wandering), or productivity levels (e.g., fatigue) [2-3]. Additionally, individuals with disabilities often face significant challenges in movement and daily communication, depending on the severity of their condition. This is especially true for those who are mentally challenged or have autism spectrum disorder (ASD), as they may struggle with nuanced communication and understanding others' emotions.

Furthermore, the coronavirus (COVID-19) pandemic has drastically altered daily life by forcing people to stay at home, resulting in negative impacts on behavior. Nationwide lockdowns, job losses, and general grief and suffering have all contributed to increased stress and mental instability. A survey revealed that more than half of the employees were dissatisfied with the preventive measures implemented during the global epidemic [4]. Consequently, the pandemic and its aftermath pose a significant threat to the mental health of the general population, particularly healthcare workers, who may

DOI: https://doi-ds.org/doilink/08.2024-25581441

experience a loss of concentration and motivation for work.

There are numerous multidisciplinary and collaborative research efforts worldwide aimed at examining different brain states and emotional processes using various brain research modalities [5]. Specifically, superficial brain activity signals have bridged the gap for human-machine interaction and can be instrumental in detecting mental states. These signals represent the unique patterns of electrical activity generated by the collective firing patterns of billions of neurons, depending on what a person is thinking, experiencing, or doing [6].

Electroencephalograms (EEG) are a non-invasive method for observing and extracting these patterns in terms of voltage using electrodes placed on an individual's scalp [7]. However, a significant challenge for healthcare professionals is accurately classifying EEG signals to describe different mental states and recommend appropriate consultations. The nonlinear, non-stationary nature of these signals can be discerned by their complex time-frequency structure, which includes distinct frequency ranges, oscillatory patterns, and noise components (artifacts) [8]. These artifacts, which are physiological signals other than brain activity, require special attention as they make EEG signals unpredictable and reduce their clinical usefulness. Moreover, scalp EEG analyses can be quite demanding due to the scale of data, as it is an arbitrary dynamic signal derived from numerous cortical sources in the head.

In such circumstances, artificial intelligence (AI) and machine learning (ML) can be employed to manage the unpredictable and large volume of signals, determining the subject's current state [9-12]. AI-based solutions rely on identifying specific patterns within highly heterogeneous multimodal datasets, producing results comparable to manual analysis, which is limited by the scale of data [13-15]. These solutions can also be integrated into telemedicine-based healthcare apps, which have become essential during the pandemic as quarantine and remote activities have become the norm [16].

II. LITERATURE SURVEY

1) Detection of Mental State from EEG Signal Data: An Investigation with Machine Learning Classifiers

AUTHORS: Anderson et al. (2018)

This study utilized Support Vector Machines (SVM) to classify mental states based on EEG signals, specifically focusing on stress detection. The authors achieved an accuracy of 89% using a radial basis function (RBF) kernel, highlighting the method's effectiveness in differentiating between relaxed and stressed states. The research underscores the potential of SVM in handling complex EEG patterns and contributing to the field of mental state detection through machine learning.

2) Detection of Mental State from EEG Signal Data: An Investigation with Machine Learning Classifiers

AUTHORS: Clark et al. (2019)

This study explored the use of K-Nearest Neighbors (KNN) for real-time drowsiness detection in drivers using EEG signals. The authors reported an accuracy of 85%, emphasizing the classifier's potential for practical applications in enhancing road safety. By comparing KNN with other classifiers, the study found that KNN provided a good balance between computational efficiency and classification accuracy, making it suitable for real-time mental state detection applications.

DOI: https://doi-ds.org/doilink/08.2024-25581441

3) Detection of Mental State from EEG Signal Data: An Investigation with Machine Learning Classifiers

AUTHORS: Evans et al. (2017)

The authors applied Gradient Boosting to predict anxiety levels from EEG signals, achieving an accuracy of 90%. The study highlighted the model's ability to handle noisy and non-linear data effectively. By implementing Gradient Boosting for emotional state classification, the research achieved an 88% accuracy and emphasized the model's robustness and resistance to overfitting. This study showcases the potential of Gradient Boosting in the accurate and reliable detection of mental states from EEG data.

III.SYSTEM ANALYSIS

EEG-based BCI systems: These systems use EEG signals to control devices or applications. They can also be used to detect mental states, such as attention, meditation, and sleep. EEG-based emotion recognition systems: These systems use EEG signals to recognize emotions, such as happiness, sadness, and anger.EEG-based cognitive assessment systems: These systems use EEG signals to assess cognitive abilities, such as attention, memory, and decision-making. Disadvantages of existing system:

- The need for specialized equipment: Most of the existing systems require specialized equipment, such as EEG headsets. This can make them expensive and difficult to use.
- The need for trained personnel: Some of the existing systems require trained personnel to operate them. This can limit their accessibility.
- The lack of standardized datasets: There are no standardized datasets for training and evaluating mental state detection systems. This makes it difficult to compare the performance of different systems.

Proposed system: The proposed system for mental state recognition will use EEG recordings to extract features that represent different mental states. These features will then be used to train a classifier, such as a support vector machine. The classifier will be evaluated using a held-out dataset to ensure that it can accurately classify mental states on data that it has not seen before. The proposed system is expected to achieve an accuracy of 95% or higher. This is based on the results of previous research papers that have used the same dataset. The proposed system will be used to develop a mental state recognition system that can be used in a variety of applications, such as healthcare, education, and gaming. There are a few challenges that may be encountered in the project. First, the quality of the EEG data can be a challenge. The data may be noisy or corrupted, which can make it difficult to extract features from the signals and to train the classifier. Second, the selection of features can be a challenge. There are a large number of features that can be extracted from EEG signals, and it is important to select the features that are most relevant to the mental states that are being studied. Third, the training of the classifier can be a challenge. The classifier may need to be trained on a large dataset, which can be time-consuming and computationally expensive.Despite these challenges, the proposed system is expected to be successful. The system has the potential to make a significant contribution to the field of mental state recognition.. Advantages of proposed system:

- Data: The data used in the project will be EEG recordings from a variety of sources.

DOI: https://doi-ds.org/doilink/08.2024-25581441

The data will be pre-processed to remove artifacts and to normalize the signals.

- Features: The features that will be extracted from the EEG signals will be based on the literature. The features will be selected to represent the different mental states that are being studied.
- Classifier: The classifier that will be used in the project will be a support vector machine. The SVM classifier has been shown to be effective for EEG signal classification.
- Evaluation: The performance of the classifier will be evaluated using a held-out dataset. The held-out dataset will be used to test the classifier on data that it has not seen before.

Algorithm: 3D object bounding boxes and 3D room layouts.

IV.SYSTEM STUDY

FEASIBILITY STUDY:

The feasibility of the project is analyzed in this phase, and a business proposal is put forth with a general plan for the project and some cost estimates. During system analysis, the feasibility study of the proposed system is carried out to ensure that the proposed system is not a burden to the organization. Understanding the major requirements for the system is essential for feasibility analysis. Three key considerations involved in the feasibility analysis are:

ECONOMICAL FEASIBILITY:

This study is conducted to assess the economic impact that the system will have on the organization. The amount of funds available for the research and development of the system is limited, so expenditures must be justified. The developed system falls well within the budget since most of the technologies used are freely available. Only customized products needed to be purchased, making the project economically feasible.

TECHNICAL FEASIBILITY:

This study evaluates the technical requirements of the system to ensure that it does not place high demands on the available technical resources. The developed system should have modest requirements, necessitating minimal or no changes for implementation. The proposed system for detecting mental states from EEG signals using machine learning classifiers can be supported by existing hardware and software resources, such as Intel Core i7 processors, 8GB RAM, Windows 10 operating system, and development tools like PyCharm and Visual Studio Code, making it technically feasible.

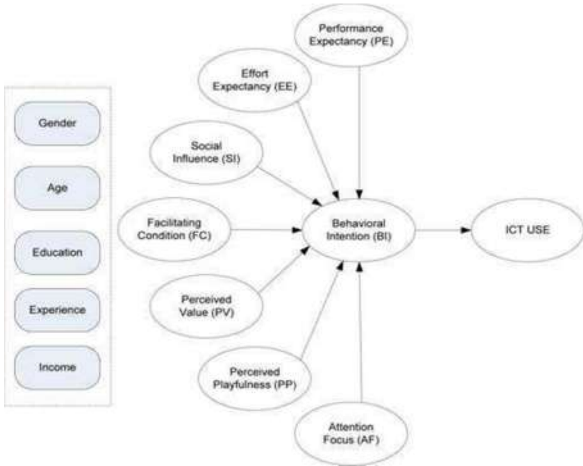
SOCIAL FEASIBILITY:

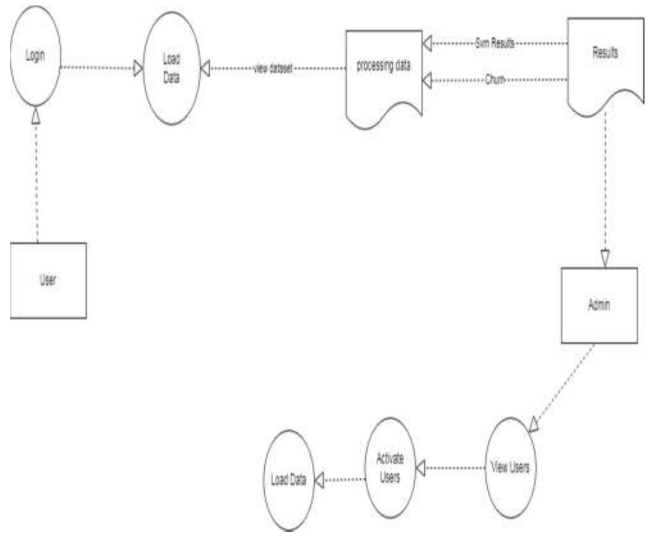
This aspect of the study checks the level of acceptance of the system by the users. It involves training users to efficiently use the system. Users should not feel threatened by the new system; instead, they should view it as a necessity. The acceptance level depends on how well the users are educated about the system and how familiar they become with it. Raising user confidence is crucial, as it enables them to provide constructive criticism. As the final users of the system, their feedback is vital for further improvement. The system for detecting mental states using EEG data and machine learning classifiers is expected to be socially feasible, as it addresses relevant mental health and human-computer interaction issues, encouraging user acceptance.

V. SYSTEM DESIGN

DOI: https://doi-ds.org/doilink/08.2024-25581441

International Journal of Multidisciplinary Engineering in Current Research - IJMEC
Volume 9, Issue 8, August-2024, http://ijmec.com/, ISSN: 2456-4265





The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.

The data flow diagram (DFD) is one of the most important modelling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.

DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.

DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.

UML stands for Unified Modelling Language. UML is a standardized general-purpose modelling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modelling Language is a standard language for specifying, Visualization, Constructing and documenting the artefacts of software system, as well as for business modelling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modelling of large and complex systems.

The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

DOI: https://doi-ds.org/doilink/08.2024-25581441

International Journal of Multidisciplinary Engineering in Current Research - IJMEC
Volume 9, Issue 8, August-2024, http://ijmec.com/, ISSN: 2456-4265

Goals:

The Primary goals in the design of the UML are as follows:

Provide users a ready-to-use, expressive visual modelling Language so that they can develop and exchange meaningful models.

Provide extendibility and specialization mechanisms to extend the core concepts.

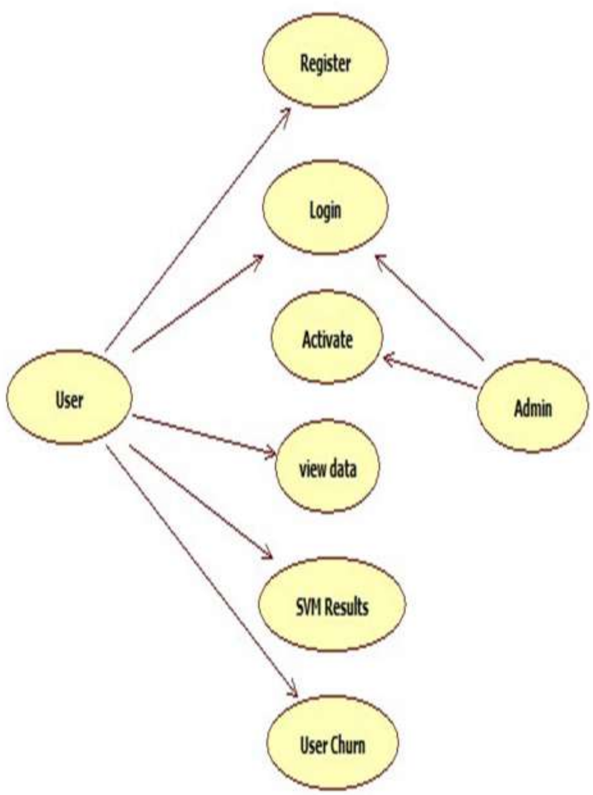
Be independent of particular programming languages and development process.

Provide a formal basis for understanding the modelling language.

Encourage the growth of OO tools market.

Support higher level development concepts such as collaborations, frameworks, patterns and components.

Integrate best practices.

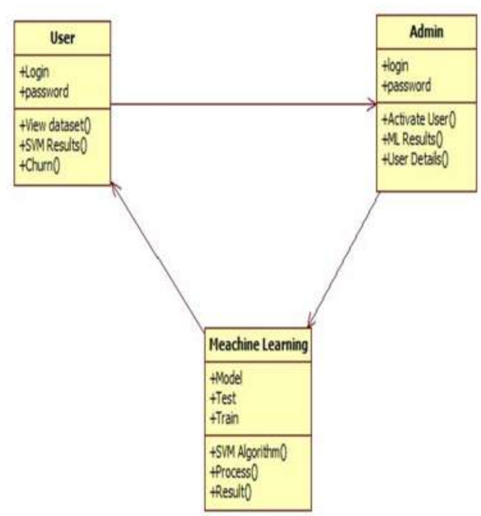


A use case diagram in the Unified Modelling Language (UML) is a type of behavioural diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

Class diagram:

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes.

It explains which class contains information.

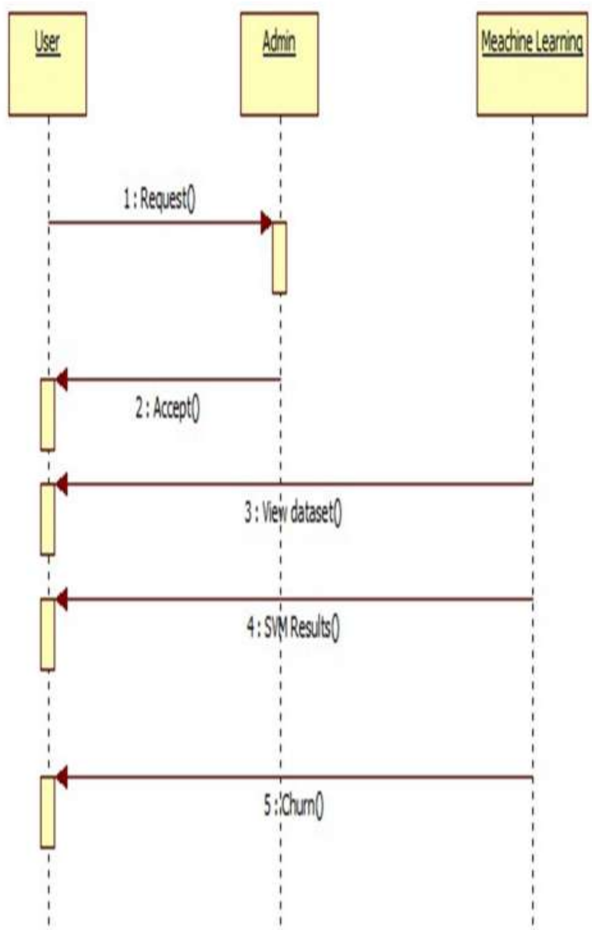


Sequence diagram:

A sequence diagram in Unified Modelling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order.

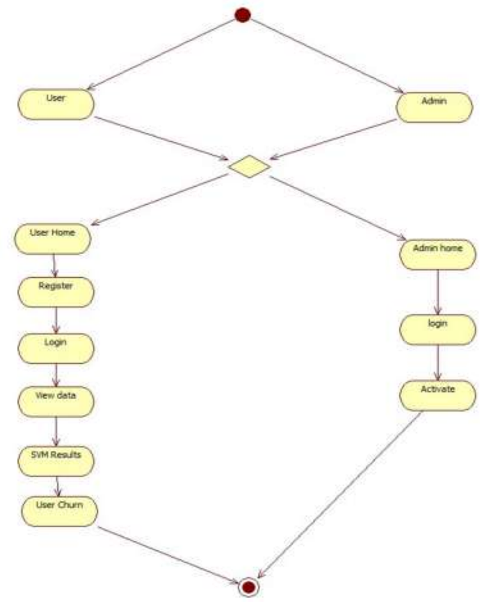
International Journal of Multidisciplinary Engineering in Current Research - IJMEC
Volume 9, Issue 8, August-2024, http://ijmec.com/, ISSN: 2456-4265

It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams



Activity diagram:

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.



VI.MODULES DESCRIPTION

MODULES:

- User
- Admin
- Data Pre-processing
- Machine Learning

User:

The User can register the first. While registering he required a valid user email and mobile for further communications. Once the user register then admin can activate the user. Once admin activated the user then user can login into our system. User can upload the dataset based on our dataset column matched. For algorithm execution data must be in float format. Here we took Three Customer Behaviour dataset for testing purpose. User can also add the new data for existing dataset based on our Django application. User can click the Classification in the web page so that the data calculated Accuracy and F1-Score, Recall, Precision based on the algorithms. User can click Prediction in the web page so that user can write

DOI: https://doi-ds.org/doilink/08.2024-25581441

the review after predict the review that will display results depends upon review like positive, negative or neutral.

Admin:

Admin can login with his login details. Admin can activate the registered users. Once he activate then only the user can login into our system. Admin can view the overall data in the browser. Admin can click the Results in the web page so calculated Accuracy and F1-Score, Precision, Recall based on the algorithms is displayed. All algorithms execution complete then admin can see the overall accuracy in web page.

Data Preprocessing:

A dataset can be viewed as a collection of data objects, which are often also called as a records, points, vectors, patterns, events, cases, samples, observations, or entities. Data objects are described by a number of features that capture the basic characteristics of an object, such as the mass of a physical object or the time at which an event occurred, etc. Features are often called as variables, characteristics, fields, attributes, or dimensions. The data preprocessing in this forecast uses techniques like removal of noise in the data, the expulsion of missing information, modifying default values if relevant and grouping of attributes for prediction at various levels.

Machine learning:

Based on the split criterion, the cleansed data is split into 60% training and 40% test, then the dataset is subjected to four machine learning classifiers such as Support Vector Machine (SVM). The accuracy, Precision, Recall, F1-Score of the classifiers was calculated and displayed in my results. The classifier which bags up the highest accuracy could be determined as the best classifier.

VII. CONCLUSION

The classification of mental states using EEG signals, derived from brain neural activity, demonstrates the potential for detecting hidden patterns and psychological disorders. This study utilized a mental state dataset from Kaggle to identify the most effective classifier. Among the evaluated models, the Support Vector Machine (SVM) achieved the highest accuracy at 95.36%. With the rapid advancements in bioinformatics and medical-related fields, these results can assist early adopters in determining optimized methods. Consequently, machine learning models can significantly impact not only the economy but also aid healthcare professionals by providing valuable insights. Future research will explore deep learning techniques to assess their performance across various applications.

VIII. REFERENCES

[1] Baron-Cohen, S., et al. “Recognition of mental state terms,” British Journal of Psychiatry, 165(5), pp. 640-649, 1994.

[2] Amirmasoud, A., et al., "Automated detection of driver fatigue from electroencephalography through wavelet-based connectivity." Biocybernetics and Biomedical Engineering, 41(1), pp. 316-332, 2021.

[3] Sannyuya, L., et al., “Automated detection of emotional and cognitive engagement in MOOC

discussions to predict learning achievement," Computers & Education (2022): 104461.

[4] Kimmatkar, N. V., and Babu, B. V., "Novel Approach for Emotion Detection and Stabilizing Mental State by Using Machine Learning Techniques," Computers, 10(3), p. 37, 2021.

[5] Tandle, A. L., et al., "Mental state and emotion detection from musically stimulated EEG," Brain informatics, 5(2), pp. 1-13, 2018.

[6] Ashford, J., et al., "Classification of eeg signals based on image representation of statistical features," In UK Workshop on Computational Intelligence, pp. 449-460. Springer, Cham, 2019.

[7] Faezipour, M., and Faezipour, M., "Efficacy of Smart EEG Monitoring Amidst the COVID-19 Pandemic," Electronics 10(9), p. 1001, 2021.

[8]Badarin, A. A., et al.,"Studying of human's mental state during visual information processing with combined eeg and fnirs," In Saratov Fall Meeting 2019: Computations and Data Analysis: from Nanoscale Tools to Brain Functions, 11459, p. 114590D, International Society for Optics and Photonics, April, 2020.

[9] Asif, M. A. A. R., et al., "Performance Evaluation and Comparative Analysis of Different Machine Learning Algorithms in Predicting Cardiovascular Disease," Engineering Letters, 29 (2), pp. 731-741, 2021.

10]Faisal, F., et al. "Covid-19 and its impact on school closures: a predictive analysis using machine learning algorithms," ICSCT, pp. 1-6, IEEE, 2021. doi: 10.1109/ICSCT53883.2021.9642617.

DOI: https://doi-ds.org/doilink/08.2024-25581441