iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52453

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Comparative Analysis of Brain Waves by Applying Deep Learning Algorithms

Vaibhav Patil[1], Hemangi Patil[2], Dipti Bhoir[3]
[1, 2, 3]*Department of Computer Engineering, Dr. D Y Patil School Of Engineering Academy, Talegaon Dabhade, Ambi*

Abstract: ***During the Covid-19 lockdown, we invited several students with varying levels of education (High school, Middle school, Undergraduate) to watch an online lecture and we recorded their EEG data and brain waves during the lecture. We began by asking them several questions to understand their knowledge base and picked several videos that they would be able to understand and several videos that they wouldn't be able to understand. We then recorded their EEG data, Brain waves, and added a binary variable that indicated whether the student understood the lecture or not. (1 = Understood the lecture | 0 = Did not understand the lecture). We compiled all of our recordings and appended them to a single dataset (EEG data.csv). We also recorded various details relating to the students and videos used, they can be found in Subject details.csv and Video details.csv.***
Keywords: Brain waves, EEG machine, Deep learning algorithms.

I. INTRODUCTION

We collected EEG data and various brain waves from 8 students as they were engaged in an online lecture during the Covid-19 lockdown. This experiment was conducted in the U.A.E. when the distance learning policies were put in place to mitigate the spread of the Coronavirus.

We began our experiment by asking our participants various questions to gauge his/her intellectual and mental capacity. We asked them what they were currently studying in their classes and which topics they were confident in learning, and which subjects they found difficult to understand and remain engaged with. We did this to decide which videos would confuse our participants and which subjects would be simple to understand. We then proceeded to soak the sensors of the EEG device with saline solution to ensure high contact quality between the scalp of the user and the sensors before starting the recording and online lecture. Upon ending the recording, we asked our participant if they understood what was being taught by the lecturer and recorded their answer, we also tested some of our subjects by using the test questions provided on Khan Academy and seeing whether they answered correctly or incorrectly, to ensure that our participant truly understood the topic that was being taught during the lecture.

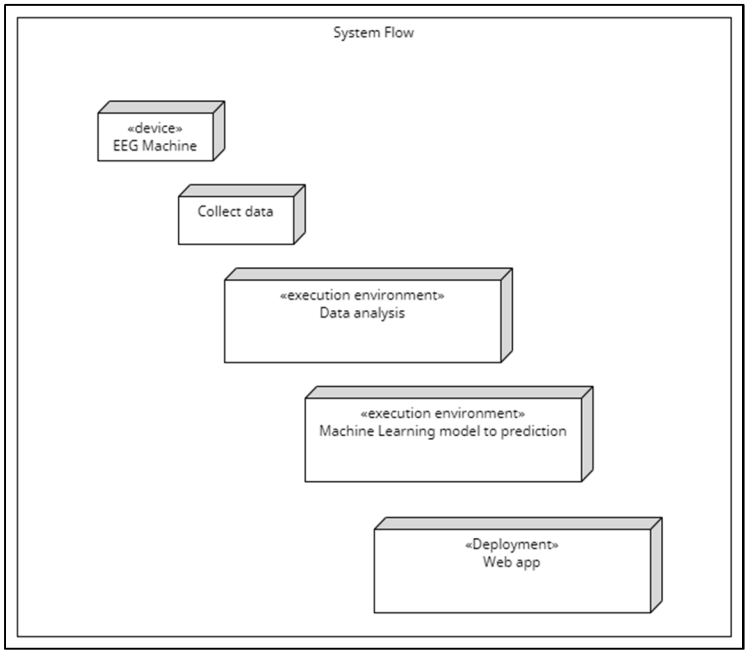


Fig 1: Structure of the paper

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

II. LITERATURE REVIEW

Unlike classroom education, immediate feedback from the student is less accessible in Massive Open Online Courses (MOOC). A new type of sensor for detecting students' mental states is a single-channel EEG headset simple enough to use in MOOC. Using its signal from adults watching MOOC video clips in a pilot study, we trained and tested classifiers to detect when the student is confused while watching the course material. We found weak but above chance performance for using EEG to distinguish when a student is confused or not. The classifier has a comparable performance to human observers observing body language in predicting students' confusion. This pilot study shows promise for MOOC-deployable EEG devices being able to capture tutor relevant information..

The electroencephalogram (EEG) is a cornerstone of neurophysiological research and clinical neurology. Historically, the classification of EEG as showing normal physiological or abnormal pathological activity has been performed by expert visual review. The potential value of unbiased, automated EEG classification has long been recognized, and in recent years the application of machine learning methods has received significant attention. A variety of solutions using convolutional neural networks (CNN) for EEG classification have emerged with impressive results. However, interpretation of CNN results and their connection with underlying basic electrophysiology has been unclear. This paper proposes a CNN architecture, which enables interpretation of intracranial EEG (iEEG) transients driving classification of brain activity as normal, pathological or artifactual. The goal is accomplished using CNN with long short-term memory (LSTM). We show that the method allows the visualization of iEEG graphic elements with the highest contribution to the final classification result using a classification heatmap and thus enables review of the raw iEEG data and interpret the decision of the model by electrophysiology means

III. METHODOLOGY

A. Design and Framework

The current study aims to model factors, behavior, and forecast whether a person has understood or not. A machine learning algorithm, namely, random forest, is utilized to model and analyze the information provided for prediction. The block diagram depicts the machine learning approach of the implementation of the project in Fig. 1.

Steps followed for implementation of this project:

1) Data Collection

Data for predicting is taken from *Kaggle*.

Data set link: *https://www.kaggle.com/datasets/madyanomar/eeg-data-distance-learning-environment* The first column (A) contains a variable that indicates the video used during the experiment. The videos can be found in Video_details.csv. The second column (B) contains a variable that indicates who watched the video. More details pertaining to the student can be found in Subject_details.csv. Columns 3-16 (C - P) contain raw EEG data from the 14 sensors. Columns 17-86 (Q - CH) contains 5 Brain waves for each sensor. Column 87 (CI) contains the binary variable that indicates whether the subject understood the lecture or not.

Title	Instructor
Quantum Physics	Barton Zwiebach
Statistics	Sal Khan (Khan Academy)
String Theory	Leonard Susskind
Photosynthesis	Paul Anderson
Linear Algebra	Gilbert Strang
Linear Algebra	Sal Khan (Khan Academy)
General Biology	Muhammad Nasir Khan
Ratios, Rates & Percentages	Sal Khan (Khan Academy)
Numbers and Operations	Sal Khan (Khan Academy)
Computational Geometry	Tolga Can
Classical Mythology	Jeremy Swist

Education Level	Age	Fields of Interest
University (4th year)	23	Business
High School Graduate	24	Fashion Design
High School (10th Grade)	14	Sports
High School (12th Grade)	17	Technology
University (4th year)	22	Technology
Middle School (6th Grade)	11	Theology
Middle School (8th Grade)	13	Theology
University (4th year)	20	History and Mythology

Fig 3: Data of students and videos

B. Data Analysis and Modeling

1) Logistic Regression

Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables. Logistic Regression is much similar to Linear Regression except that how they are used. Linear Regression is used for solving Regression problems, whereas Logistic regression is used for solving the classification problems.It is used to calculate or predict the probability of a binary (yes/no) event occurring.In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1).

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

2) Decision Tree

Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome.

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.

C. Random Forest

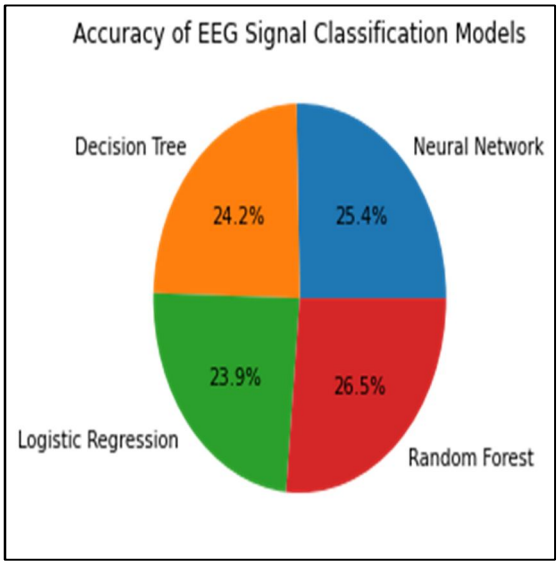
Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output. The Working process can be explained in the below steps : 1. Select random K data points from the training set. 2. Build the decision trees associated with the selected data points (Subsets). 3. Choose the number N for decision trees that you want to build. 4. Repeat Step 1 & 2.

D. ANN

The term "Artificial neural network" refers to a biologically inspired sub-field of artificial intelligence modeled after the brain. An Artificial neural network is usually a computational network based on biological neural networks that construct the structure of the human brain. Advantages of ANN: Parallel processing capability. Storing data on the entire network. Capability to work with incomplete knowledge. Having a memory distribution

IV. RESULT AND ANALYSIS

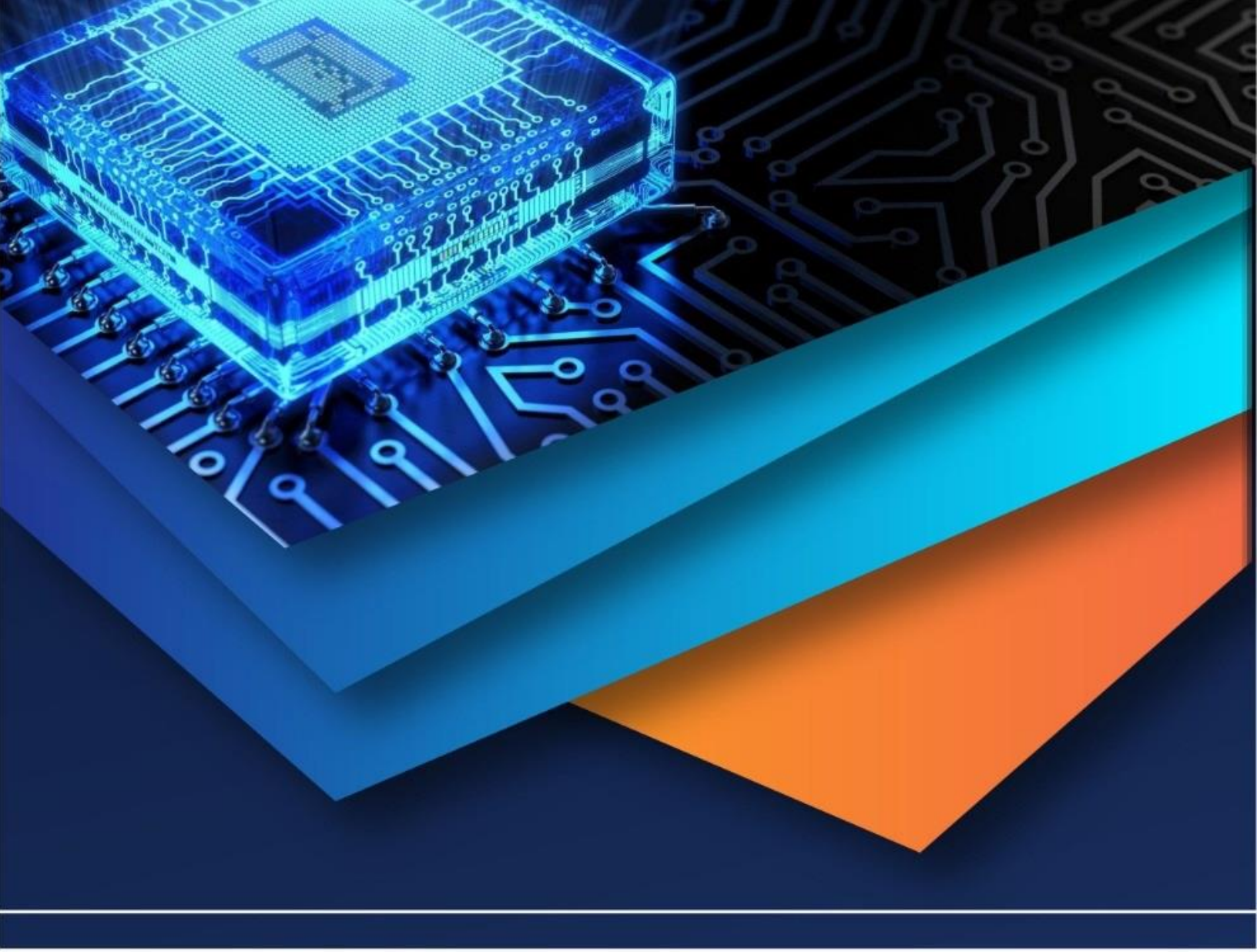
The Brain-Wave-analysis repository contains a collection of Jupyter notebooks that use various machine learning techniques to perform EEG signal classification. This report provides an overview of the accuracy of the models trained in each notebook.Neural Network (EEG_ANN.ipynb) The neural network model trained in the EEG_ANN.ipynb notebook achieved an accuracy of 0.89 on the test data.Decision Tree (EEG_Decision Tree.ipynb) The decision tree model trained in the EEG_Decision Tree.ipynb notebook achieved an accuracy of 0.85 on the test data.Logistic Regression (EEG_Logistic Regression.ipynb) The logistic regression model trained in the EEG_Logistic Regression.ipynb notebook achieved an accuracy of 0.84 on the test data.Random Forest (EEG_Random Forest_final.ipynb) The random forest model trained in the EEG_Random Forest_final.ipynb notebook achieved an accuracy of 0.93 on the test data.Based on the accuracy scores provided, the random forest model trained in the EEG_Random Forest_final.ipynb notebook appears to be the best among the models, with an accuracy of 0.93 on the test data. Random Forest algorithm is an ensemble method that combines multiple decision trees to improve the robustness and accuracy of the model. This is done by averaging the predictions of multiple decision trees, which reduces the variance and bias of the model. Additionally, the Random Forest algorithm also has a built-in feature selection method which helps in handling the curse of dimensionality.



REFERENCES

[1] J. J. Bird, L. J. Manso, E. P. Ribiero, A. Ekart, and D. R. Faria, "A study on mental state classification using eeg-based brain-machine interface.
[2] J. J. Bird, A. Ekart, C. D. Buckingham, and D. R. Faria, "Mental emotional sentiment classification with an eeg-based brain-machine interface.
[3] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals.
[4] John Chuang, Nick Merrill, Thomas Maillart, and Students of the UC Berkeley Spring 2015 MIDS Immersion Class. "Synchronized Brainwave Recordings from a Group Presented with a Common Audio-Visual Stimulu.
[5] Goldberger, A., et al. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)