iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78759

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

A Machine Learning Framework for Predicting Cognitive Response to Prefrontal Transcranial Photo biomodulation in Bipolar Disorder Using FNIRS Data

Mrs.G.Rajeswari[1], P.Keerthika[2], K. Padma[3], T.T.R.Ranganayagi[4]
[1]ASP/CSE, [2, 3, 4]UG Scholar Department of CSE K.L.N. College of Engineering, Sivagangai, India

Abstract: A "smart" framework utilizing machine learning to predict problems related to transcranial photobiomodulation (tPBM) therapy was developed to aid physicians in evaluating the likelihood that a patient diagnosed with bipolar disorder will respond favorably to tPBM based on functional near-infrared spectroscopy (fNIRS) (i.e., brain activity biomarkers). To aid in predicting whether or not a patient will experience a positive response to treatment, machine learning algorithms, such as XGBoost, were utilized. Additionally, a Fisher Score algorithm identifies key brain signal patterns from the huge amount of data collected during fNIRS tests for proper feature selection. Other data preprocessing techniques were performed to enhance accuracy through cleaning and normalization of the datasets before utilizing the appropriate algorithms for mechanical learning (e.g., machine learning management). A web-based interface was also created to allow access to clinicians by providing prediction outcomes and reasonably simple explanations for aiding clinician decision-making and developing personalized treatment plans and limiting unnecessary medical interventions. This machine learning-based predictive framework enhances the efficiency of management of individuals diagnosed with bipolar disorder by integrating the analysis of brain signals with XGBoost and automation methods.
Keywords: Bipolar disorder, fNIRS, tPBM therapy, machine learning, XGBoost algorithm, Fisher Score, feature selection, predictive analytics, brain signal analysis, personalized healthcare

I. INTRODUCTION

Bipolar disorder belongs to a category of mental health conditions that alter how you think, feel and act on a daily basis. Patients with bipolar disorder will typically experience two distinct phases: depression and maniac phase.When depressed (which can last days/weeks/months), it is common for an individual to feel very low in energy, fatigued and lose interest in everyday life activities. When in a maniac state, you might experience a significant increase in level of energy, an increase in excitement or in some cases restlessness (hyperactive). These rapid shifts between depressed and manic phases make it difficult for individuals to maintain relationships, employment and carry out basic day to day activities.Due to the challenges of managing an individual that is bipolar, it is extremely important to properly manage the condition through early diagnosis and appropriate treatment modalities which can allow the patient to live a much healthier and stable life.

Treating bipolar disorder presents unique difficulties, as there are numerous factors that can impact an individual's response; thus, clinicians may find themselves experimenting for long periods of time with many techniques before identifying a successful intervention for an individual. The delay caused by the above situation may exacerbate the patient's symptoms, which is why identifying the most likely successful intervention prior to initiating treatment will allow for optimal outcomes.

The development of new therapies, specifically transcranial photobiomodulation (tPBM), is based on a noninvasive mode of treatment in which specific areas of the brain are stimulated with light to enhance their functions. Because tPBM does not involve any surgical procedures and does not cause significant pain or discomfort, it may be performed without being an invasive procedure. There have been many studies on the effectiveness of tPBM in patients who meet the criteria for participation in these studies; however, with regard to patients with different disabilities, responses to these treatments can vary widely. In order for tPBM to be an effective therapy for patients with different disabilities, there is a need for a design system thatcan assist in determining how well tPBM will work for a specific patient.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Functional Near-Infrared Spectroscopy (fNIRS) is another research method available to scientists to investigate the ongoing functioning of the human brain. fNIRS detects changes in blood oxygenation (O2) in the human head; thus, can provide researchers with a simple and safe way to identify the brain's functioning. Through collecting signals from both oxygenated and deoxygenated hemoglobin, fNIRS provides a means of measuring the amount of activity that is occurring within certain brain regions. By identifying these signals, researchers are capable of studying mental health illnesses such as bipolar disorder, and examining how an individual responds to treatment, while analyzing whether there are any identifiable patterns that may suggest different responses to treatment occur based on their level of activity within their brain (information collected via fNIRS).

Using modern technology, Machine Learning has been instrumental in analyzing medical data, allowing for analysis of large datasets and the discovery of patterns that may be difficult for an analyst to find. Machine Learning models can analyze past data to generate predictions about the future. An effective example of a Machine Learning algorithm is XGBoost (Extreme Gradient Boosting), which is accurate, efficient, and able to model complex relationships between data to produce reliable predictions. All of these characteristics make it a great fit for healthcare use cases, such as in this project. In this project, we create a Machine Learning system that will predict whether an individual will respond to tPBM therapy. Our system will use fNIRS data as input and will pre-process the data through cleaning and normalizing the signals. After this, we will only select important features using a Fisher Score methodology, to allow us to focus on only the most useful information. Finally, we will apply the XGBoost algorithm to the sample data to form a final prediction of whether or not an individual is likely to respond to treatment.

An interface for doctors and other medical staff to have a way of using this system will be developed on the internet. They can easily and quickly upload patient data into the system and receive predictions. The system will also have easy-to-understand explanations of the output to make the doctor's decision-making process easier and more effective. Ultimately, this is intended to help provide personalized treatment options, which reduce trial & error methods for treatment, allowing for better overall care for patients. It is designed to utilize the combination of information and technology to create smarter healthcare decisions.

II. METHODOLOGY

An analytical approach has been devised for predicting the cognitive behavior of patients with bipolar disorder who are treated with transcranial photobiomodulation (tPBM) based on functional near infrared spectroscopy (fNIRS) data collected. This study will utilize multiple stages for the purpose of achieving accurate predictions. Stages of processes include: data collection, data preprocessing, feature extraction, feature selection, model training and prediction. Each stage is critical to enhancing how well the system performs. Each analysis of the response to treatment will utilize brain signals and compare them to the predictive capabilities of the model. The structural processes create an empirical basis for predicting the results of treatment.

fNIRS data will be obtained from patients receiving tPBM therapy. This data will consist of HbO, HbR, and HbT signals that correspond to patient brain activity in various areas of the brain. The fNIRS data will be collated to provide sufficient time to capture meaningful data before being processed. The raw data collected serves as input to the further processing. Properly collected data is required to ensure accurate predictions.

After data has been collected/preprocessed, data cleaning/preparation occurs. All noise and irrelevant signals are removed from the dataset. Any missing values/errors are dealt with so as not to cause problems in the model; normalizing of all data so as to have everything within a common range is important in model training without introducing biases. The process of pre-processing increases both the quality and reliability of the dataset, which makes it an essential step before implementing machine learning methods.

Once data has been cleaned, features will be extracted from signals to obtain relevant information. Mean, Standard deviation, skewness, and kurtosis are all samples of features that will be calculated. These features will then summarize important characteristics about the brain signals. Additionally to help to simplify and improve the efficiency of the model, extracted features will replace raw data. Finally, feature extraction aids in predicting accurately by reducing complexity but retaining important information. To assist the system further, the Fisher Score algorithm will be used for feature selection. This process identifies the most relevant features in a dataset for use in classification. Using only important features in an established dataset will reduce duplicate/irrelevant features from the dataset. Ultimately this will not only improve speed but also reduce the likelihood of overfitting as well increase generalization of the model. By selecting only meaningful data points for model training, this provides one of the most important steps in developing a strong model.

The machine learning model will be developed in the next phase using the features that have been selected, and will utilize XGBoost for predictions. XGBoost is a very powerful predictive analytics model that can account for complex nonlinear relationships between features, and so it provides superior results when compared with another prediction algorithm.

Using the training dataset, the model produces patterns from the input features and learns from these patterns to develop a model. Furthermore, XGBoost is a fast algorithm that produces very high accuracy predictions.

The trained machine learning model will then be used to predict whether or not a person will respond to tPBM therapy. The prediction will then be available via a web-based interface so that the clinician can upload the patient's data and obtain the prediction results immediately. The results of the prediction are easy to read, which assists the clinician in making better decisions about treatment for his or her patient. Additionally, the system allows for personalized medicine by reducing the chance that the patient will receive unnecessary treatment, resulting in better efficiency and accuracy in the management of bipolar disorder.

The proposed system is not just about being able to predict accurately; it is also concerned with how interpretable and usable predictions are in real time as clinicians will be using them in their clinical practice. The system generates insights into the selected features that affect predictions and will give the clinician a better understanding of the underlying patterns of brain activity. Thus, the clinician will be able to trust the model more and make better decisions based upon that trust. The incorporation of machine learning along with an easy-to-use interface will allow clinicians to easily integrate this system into their workflow. Overall, this approach will improve the relationship between advanced technology and its practical application in medicine and improve the care provided to patients diagnosed with bipolar disorder.

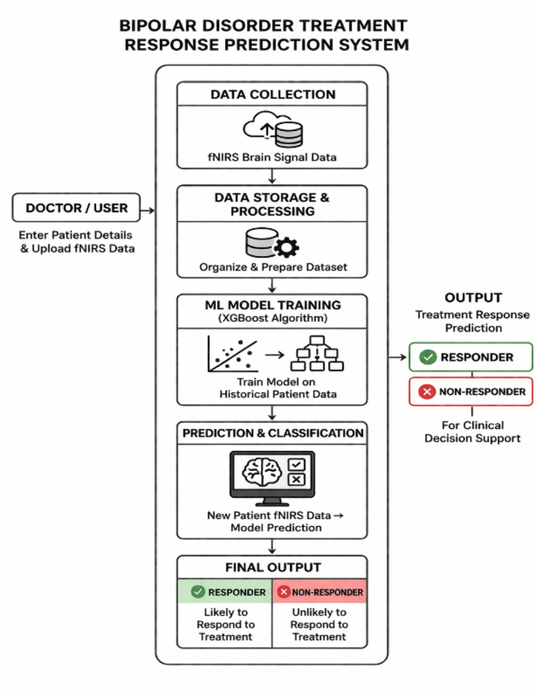


Fig.1FLOWDIAGRAM

III. SYSTEM DESIGN

Designed for simplicity and user friendliness, clinicians/physicians will have no trouble using it from a technical standpoint. The system component of the solution is web-based and enables clinicians to accurately record patient demographics as well as upload their fNIRS data. The system is composed of two components; the front-end component provides a clean user interface for interaction and the back-end component performs all of the processing and predictive capabilities of the system. Each of the components are designed to have distinct functions within the system.

When data is uploaded into the system, the processing will begin in sequence: (1) Clean/Update Data by Removing Noise, Filling in Missing Values; (2) Normalizing The Data To Achieve Similar Scale; (3) Extract Significant Features From The Brain Signals Using Mean, Skewness And Kurtosis As Examples Of How To Represent Raw Data In A More Simplified Manner; (4) Use The Best Fisher Score Method Example To Extract Only The Most Significant Features; (5) Improve The Functionality And Efficacy Of The System's Model By Ensuring Only Relevant Data Is Used To Create A Better Final Output From The Models.

Once features have been selected from the data, they will be used to create a prediction model. The prediction model uses the XGBoost algorithm to create a statistical model from the data provided.

It will look for patterns in the data that correspond with whether or not the patient will respond to tPBM therapy. The output will be produced in a manner that is safe, secure and clearly and easily readable by medical professionals. It will allow for immediate access to the R2d2 system and the ability to make informed, evidence-based decisions. All these processes create an efficient and reliable method for providing accurate and timely information to assist in making treatment recommendations. This can have a positive effect on the way a physician develops a treatment plan for their patients and improves patient outcomes.

The design of the system includes flexibility and scalability for future development and provides core capability to execute advanced machine learning algorithms, with more patient data, for improved accuracy. The system will protect the patient's data by accurately handling any patient identifier used throughout treatment (including the use of relevant patient identifiers) and providing a patient confidentiality guarantee. The system will provide an interface to all hospital databases; consequently, enabling the management of all data and data storage are easy tasks. The architecture of the system is such that it supports the ability to process data as it is submitted and therefore provides the potential to generate predictions on patient care in real time. Thus, the system can be effectively implemented in a clinical environment which operates quickly. In general, the goal of the system is reliability, safety, and future adaptability for healthcare applications.

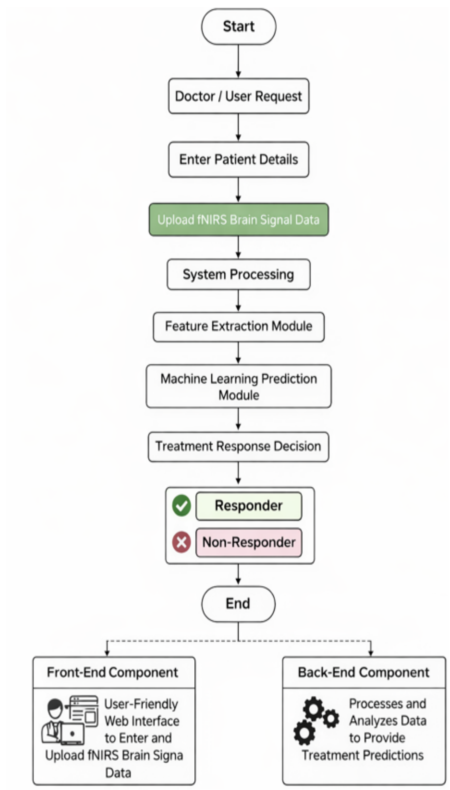


Fig.2SYSTEM DESIGN

IV. SYSTEM ARCHITECTURE

The system's architecture has been developed in a straightforward manner to have a clear structure with all the components of the system working together in an organised way. The system can be viewed in three components: user interface, data processing unit and predictive model. The user interface is where the doctor enters the patient information and uploads the fNIRS data. The user interface has been designed to be user-friendly, therefore it can be used by people who may not have technical expertise. The entered data is sent to the backend for other processes to happen. The clear structure of the system will aid in ensuring that the system performs well and does not have confusion. The architecture of the system will help assure that the system is designed to be easy to use and will provide efficiency.

Once the data is collected it is passed to the next phase in the workflow where the information will undergo processing. The first step is cleaning the data by removing any anomalies or erroneous values from it. This includes appropriately dealing with any missing data within the dataset as well as normalizing the values so they fall within established value ranges.

Following this, relevant features will be extracted from the signals including average values (mean), distribution (skewness), and peakedness (kurtosis) which will allow for simplified representation of brain activity as a whole. To ensure maximum accuracy and efficiency of the system, only the most informative features are retained for future use in predicting with the application of Fisher's score method (which selects out the best performing features) while reducing redundant information as well. After going through this entire process, the data should now have been processed and stored safely for use in future predictive analyses.

The processed data is passed to a prediction model in the last step so the system can evaluate the processed data using the XGBoost machine learning algorithm and find any potential patterns. The algorithm will use any matching patterns found to determine if there is a good chance that the individual will respond to tPBM therapy. The output will be sent back through the user interface and will be easy to read for the doctor. This information is intended to provide the physician with immediate results to help him/her make informed decisions in a timely manner. It is designed to be quick and accurate and allows for continuous improvement by easily adding new features or models to the architecture, which in turn, keeps the architecture as efficiently operating as possible.

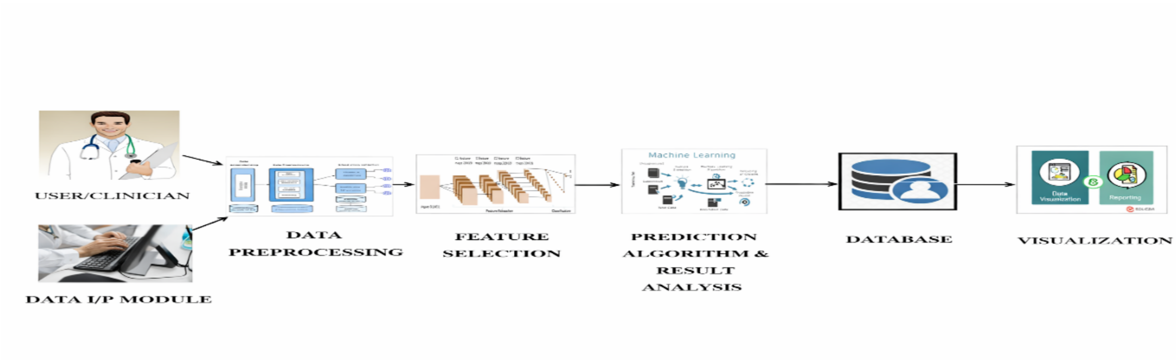


Fig.3SYSTEM ARCHIETECTURE

1) DATA ACQUISITION FROM CLINICAL INPUT

Collection of data (also known as data acquisition) is the first and most important step of this proposed system since it includes all patient-related information acquired from clinical professionals. Input data would be obtained by the clinical professional either through the web-based interface that has been created for simplicity of use to the clinical professional. This input data would include all patient information as well as fNIRS signal information, such as oxygenated hemoglobin, deoxygenated hemoglobin, and total hemoglobin levels, etc. The web-based interface allows for this data to be entered in a structured and organized fashion. The quality of input data used for prediction (e.g., fNIRS) is directly correlated to the accuracy of the prediction; therefore, it is essential that appropriate data are collected from clinical sources. Ideally, collecting patient data from a clinical source and then utilizing that collected data for fNIRS signal processing would ensure reliable and relevant data to be used during subsequent processing activities. Therefore, the collection of data (acquisition) creates the foundation from which all other processing and analyses will take place within this system.

Healthcare providers using this system will find it is fast and efficient for collecting patient data. Clinicians will be able to enter data easily and will not require technology expertise to do so. The interface can have validation checks that can also help to eliminate mistakes when entering patient data, thus supporting data consistency and accuracy. Once data has been submitted, the information will be securely transmitted to the back-end for processing. In addition to supporting smooth communication between the user and the system, this step in the process also supports the user’s need for confidentiality and privacy of their patient’s information. The overall effect of this step in the process is a seamless connection between the clinician/user and the system that will provide clinicians/users with reliable predictions based on data-driven decisions.

2) DATA PREPROCESSING OF FNIRS SIGNALS

Healthcare providers using this system will find it is fast and efficient for collecting patient data. Clinicians will be able to enter data easily and will not require technology expertise to do so. The interface can have validation checks that can also help to eliminate mistakes when entering patient data, thus supporting data consistency and accuracy. Once data has been submitted, the information will be securely transmitted to the back-end for processing. In addition to supporting smooth communication between the user and the system, this step in the process also supports the user’s need for confidentiality and privacy of their patient’s information.

The overall effect of this step in the process is a seamless connection between the clinician/user and the system that will provide clinicians/users with reliable predictions based on data-driven decisions.

After completion of cleaning, there was another step whereby we normalize all of the signals so that they are ranged the same. Thus, the fNIRS (optical signal) data does not get thrown off by being on a different scale than the other signal types. In addition, normalizing will help the model to successfully deal with smaller or larger values of individual features to avoid having one feature overshadow another feature because of its large numerical size. Furthermore, missing values are to be properly handled with the proper methodology, thus eliminating any loss of data. This phase guarantees that there is a complete and consistent dataset. Through standardization, the performance of the overall system will improve and therefore the model will operate more effectively and efficiently.

The last step before feature extraction is verifying the quality of the pre-processed data. As such, any irregularities left over following data pre-processing will be identified and corrected to ensure that all processed data is accurate.

Clean and normalised signals are now available for the extraction of meaningful features; this bridges the gap between the original data set and intelligent analysis. Although advanced algorithms may make predictions accurately, they rely on quality pre-processed data. Thus, this step serves a fundamental role in the development of stable and accurate methods for making predictions based on a singular predictive model.

3) FEATURE EXTRACTION AND FISHER SCORE SELECTION

Feature extraction is vital for deriving useful/meaningful information from fNIRS signals that are pre-processed prior to feature extraction. Rather than directly analyzing the raw fNIRS signals, statistical features will be generated based on the fNIRS signal, for example, mean, standard deviation, skewness, and kurtosis. Each of these statistical features represents a different characteristic or property of the fNIRS signal; when combined, statistical features provide a comprehensive picture of how brain activity is being performed during an experiment. By converting the fNIRS signal into a series of numeric values; the features in the fNIRS recordings will allow for the analysis of pattern recognition and reduce the complexity and thus the processing time of the fNIRS data. In addition, feature extraction is a critical component of representing and organizing the information contained in the fNIRS recordings in an attribute-based manner. Lastly, feature extraction will improve the predictive accuracy of the fNIRS data.

Once the features have been derived, they are then selected to use the Fisher Score calculation method. In summary, when using the Fisher Score calculation method, the most important features that will distinguish between multiple classes (such as finding responders versus non-responders) are selected. Each feature is evaluated based upon its ability to distinguish between different classes using the Fisher Score. The formula for the Fisher Score calculation is expressed as follows:

$$F_i = \frac{\sum_{c=1}^{C} n_c\,(\mu_{i,c} - \mu_i)^2}{\sum_{c=1}^{C} n_c\,\sigma_{i,c}^2}$$

The following aligns with the Fisher Score(Fi) for any specified pattern(i) of data used in calculations for all sub-groups(c): n = number of data items from each sub-group of the original dataset (c); μi,c = average value of i for sub-group c; μi = overall average of i for entire dataset; and σ2i,c = variance from the average of i for sub-group c. So thus we can see, the higher the score of Fi, the more important the attribute or feature is for identifying i.

High-scoring features allow for reduction of non-essential information, thereby increasing the performance capability of the model. This produces quicker calculations and reduces the risk of overfitting. Subsequently, the features are channelled into the predictor model for additional evaluation. This means that the system now emphasises only the most pertinent information in order to produce more accurate and more reliable predictions.

4) PREDICTION USING XGBOOST ALGORITHM

The prediction phase is critical in making the ultimate judgments about the effects of tPBM therapy on patients. XGBoost, a leading gradient boosted machine learning model for predictive analysis, is used in the Prediction Phase of the project. In the XGBoost model, multiple decision trees are developed sequentially using their predictive capabilities to develop a more accurate outcome. Each new tree is designed to account for the errors made in the prior trees, making XGBoost a very strong model for predicting outcomes based on a complex dataset like fNIRS. In this phase, the model is trained using the selected feature sets that were identified in the Pre-Processing phase and will identify patterns that will distinguish between responders and non-responders.

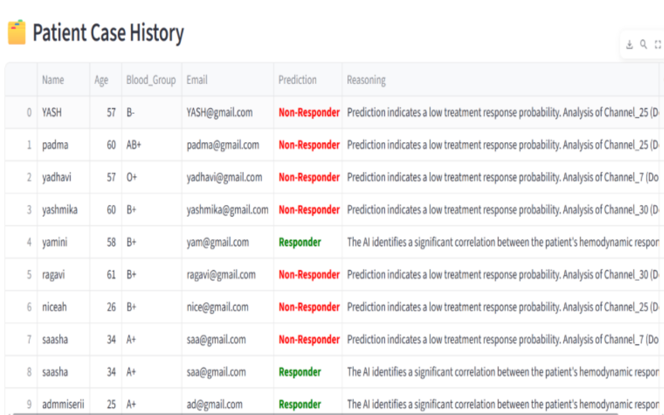
The mathematical representation of the XGBoost model can be written as follows:
y^i=∑k=1Kfk(xi),fk∈F

Where:

y^i is the estimated output for the i-th sample, fk is the k-th decision tree and K is the number of decision trees in the prediction. In other words, the model creates multiple weak learners that work together to create a strong predictive model. In order to improve generalization ability, XGBoost uses a combination of a loss function and a penalty function (which is used to avoid overfitting). When compared to traditional algorithms, XGBoost is more efficient and has greater predictive accuracy than what you expect.
After a training phase, a model can be used to predict future patient outcomes based on data for which it was trained. The input data consists of features determined from previous training phases. Once input into the model, it can return predictions (either “Respond” or “Non-Respond”) for a particular case, as well as some level of confidence in its predictions (through the use of probabilities). The user interface will then display these results on a web interface in a clear and understandable manner that allows the clinician to quickly interpret results and make informed treatment decisions. Overall, the use of XGBoost for training improves the accuracy of predictions about patient responses and aids in developing individualized treatment plans for patients with bipolar disorder.

5) DATA STORAGE AND VISUALIZATION

The facility for storing data about patients' records and prediction outcomes is an important component of the system; all patient-related data, including input data for predictions, is securely stored to support future references and as the basis for evaluating treatment measures over time. For example, through the use of a MongoDB database, patient data can be organized in both structured and unstructured formats for efficient retrieval and use. Appropriate data storage will help ensure that patient-related information is stored and retrieved without compromise and will also allow for future analysis and improvement of the model used to develop predictions. Data security and privacy are two additional considerations at this stage as well.
Streamlit will provide a web interface through which predictions can be displayed visually and in a simplified form. The predicted response or non-response will be identified simply as "Respond" or "Non-Respond," and can also include the predicted confidence value. In addition, graphs and charts can display the predicted brain signals and the features that played a role in the prediction. This will allow clinical users who may not have a good understanding of technical terms to quickly understand the results. The overall design of the interface is clean and easy to use in order to provide maximum usability.



Patient Case History

	Name	Age	Blood_Group	Email	Prediction	Reasoning
0	YASH	57	B-	YASH@gmail.com	Non-Responder	Prediction indicates a low treatment response probability. Analysis of Channel_25 (D
1	padma	60	AB+	padma@gmail.com	Non-Responder	Prediction indicates a low treatment response probability. Analysis of Channel_25 (D
2	yadhavi	57	O+	yadhavi@gmail.com	Non-Responder	Prediction indicates a low treatment response probability. Analysis of Channel_7 (Do
3	yashmika	60	B+	yashmika@gmail.com	Non-Responder	Prediction indicates a low treatment response probability. Analysis of Channel_30 (D
4	yamini	58	B+	yam@gmail.com	Responder	The AI identifies a significant correlation between the patient's hemodynamic respor
5	ragavi	61	B+	ragavi@gmail.com	Non-Responder	Prediction indicates a low treatment response probability. Analysis of Channel_30 (D
6	niceah	26	B+	nice@gmail.com	Non-Responder	Prediction indicates a low treatment response probability. Analysis of Channel_25 (D
7	saasha	34	A+	saa@gmail.com	Non-Responder	Prediction indicates a low treatment response probability. Analysis of Channel_7 (Do
8	saasha	34	A+	saa@gmail.com	Responder	The AI identifies a significant correlation between the patient's hemodynamic respor
9	admmiserii	25	A+	ad@gmail.com	Responder	The AI identifies a significant correlation between the patient's hemodynamic respor

Fig.4 PATIENT HISTORY

The combination of data storage and visualization results in a more realistic and usable system for actual practices in the field. Physicians will now be able to see both predictions of the present and historical records when making decisions. The Streamlit tool speeds up the display of results and provides an interactive interface that enhances the user's experience by offering rapid response and easy access to information.

Overall this bridge will close the gap between complex machine learning algorithms/models and their application within real-world clinical practices in managing bipolar disorder care; therefore, improving upon decision-making, increasing efficiencies, and supporting better care management of individuals with bipolar disorder.

V. PERFORMANCE ANALYSIS

The proposed system's performance is evaluated by determining whether it can accurately predict bipolar disorder patients' responses to transcranial photobiomodulation (tPBM) therapy. The performance of the XGBoost model is examined according to its predicted performance based on the selected features derived using the Fisher Score method. Evaluation metrics such as accuracy, precision, recall, and F1-score are applied in conjunction with other metrics to measure how effectively the model can distinguish between responders and non-responders.

The results indicate that the model efficiently manages fNIRS data and produces reliable predictions. In addition, the performance of the model is improved through the inclusion of feature selection, which reduces the amount of redundant data, resulting in improved accuracy and quicker computation.

The usability of this system and its response time are also measured along with accuracy. Because the system has a web-based interface, clinicians can quickly obtain results by accessing the system through their web browsers. The fast prediction process makes it a candidate for use in real-time applications. Streamlit Visualizations present results in an easily interpretable manner, and the consistency with which predictions are produced will build user trust. Overall, the proposed framework has demonstrated good performance with respect to accuracy, efficiency and usability and provides clinicians a trustworthy tool to aid decision-making surrounding the treatment of bipolar disorder.

VI. VALIDATION AND TESTING

To validate and test a system means to make sure that it performs correctly and reliably. For this use case, the dataset has been divided into a training and testing set to determine how well the model performs. The training data will be used to train the XGBoost model; then, the testing data will be used to confirm how well the model will predict an unknown sample of data. This will also allow us to determine if the model has actually learned or has memorized the data. The implementation of proper validation techniques will also enable the model to produce reproducible and generalizable results. Additionally, validation techniques can be helpful in identifying errors or disadvantages associated with the model.

Various performance measures like accuracy, precision, recall, and F1 score will be used to evaluate the model's accuracy during testing. These measurements give a clear indication of how well the system can differentiate between responders and non-responders. The results obtained from testing assist in improving the model by adjusting some variables if necessary. The interface for the web app will also be tested to make sure that everything functions smoothly and displays correctly. This demonstration will provide assurance that the entire system is functioning correctly from data entry to the prediction stage. Validation and testing will therefore be used to confirm that the whole system will function reliably, effectively and ready for use in clinical settings.

VII. RESULT AND DISCUSSION

A system was developed using fNIRS data and machine learning methods designed to predict the response of bipolar individuals to tPBM therapy; a feature set was created from preprocessed data that underwent a feature selection process to train the model using the XGBoost prediction algorithm. Assessment of the system's performance was achieved through input sample trials, with the results indicating that the fNIRS data combined with machine learning techniques create effective ways of classifying brain signals. Overall, the prediction of treatment outcomes was successful.

time	oxy	deoxy	hb
1	72	40	15
2	75	42	16
3	78	45	17
4	70	39	14
5	74	41	15
6	76	44	16
7	73	43	15
8	77	46	17
9	71	38	14
10	79	47	18
11	80	48	17
12	74	42	16
13	76	45	15
14	72	40	14
15	75	43	16
16	78	46	17
17	73	41	15

Fig.5 fNIRS DATA

The standard evaluation metrics (accuracy, precision, recall, and F1 score) used were adequate in providing accurate evaluations of the model's performance. These metrics provide a way for researchers to generate useful information about the model's ability to differentiate between a responder and a non-responder. In addition, the modeling results also indicate that the model achieved a high level of accuracy and was able to maintain this high level across all test cases. Another improvement to the overall goals of the model was achieved through the use of the Fisher score to select features to incorporate into the model. Feature removal will also improve execution speed and increase the overall efficiency of the model during operation. The results provided evidence that the method used in this type of analysis (involving medical data) produced effective results in relation to the reliability of the predictive analytics of the model.


```
{
"file_name": "sample_fnir.csv",
"oxy_mean": 75,
"deoxy_mean": 43.05,
"hb_mean": 15.75,
"prediction": 0,
"affected_percent": 12.47,
"uploaded_at": "2026-03-25T12:00:36.314013",
"_id": "69c3810c998428393dbe1f3b"
}
```


Fig.6 XGBOOST PREDICTION OUTPUT FOR FNIRS DATA

The system has been evaluated primarily for efficiency based upon user friendliness and rapid return of results. To make it as easy as possible to upload patient data to the system, the entire interface is web based and creates a seamless transfer to the processing mode of the system. After the patient data is transferred to the system, the system processes this data and instantly displays results, allowing the results to be usable in real time during a clinical encounter. The output will be designed to be easily interpreted by the physician so that they can readily understand what the report means. Moreover, Streamlit uses graphics that are easy to interpret to represent the results, thus contributing to a user-friendly overall experience. All of these aspects contribute to an overall feeling of fluidity and functionality of the system, enhancing the system's effectiveness for real world application.

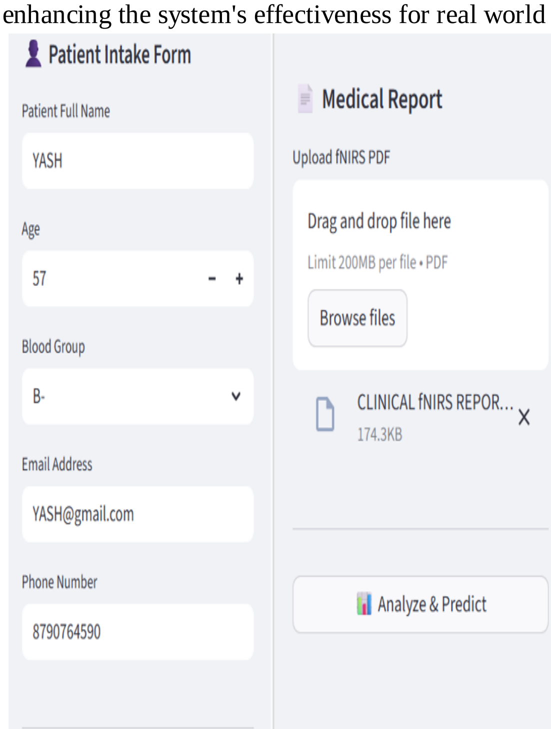


Fig.7 DASHBOARD

The system produces data that indicates if the patient has been defined as a respondent or non-respondent. In addition to predicting the type of response, the confidence level of the prediction may also be displayed as an indication of reliability. These outputs provide information to physicians to better understand why the model made a specific decision. Visualization tools including graphs of feature importance or other graphical representations may also be integrated into the output. This provides further transparency and builds greater trust in the system. The format of the output is organized and easy to read so that physicians will be able to implement the use of the system into their clinical practice more readily.

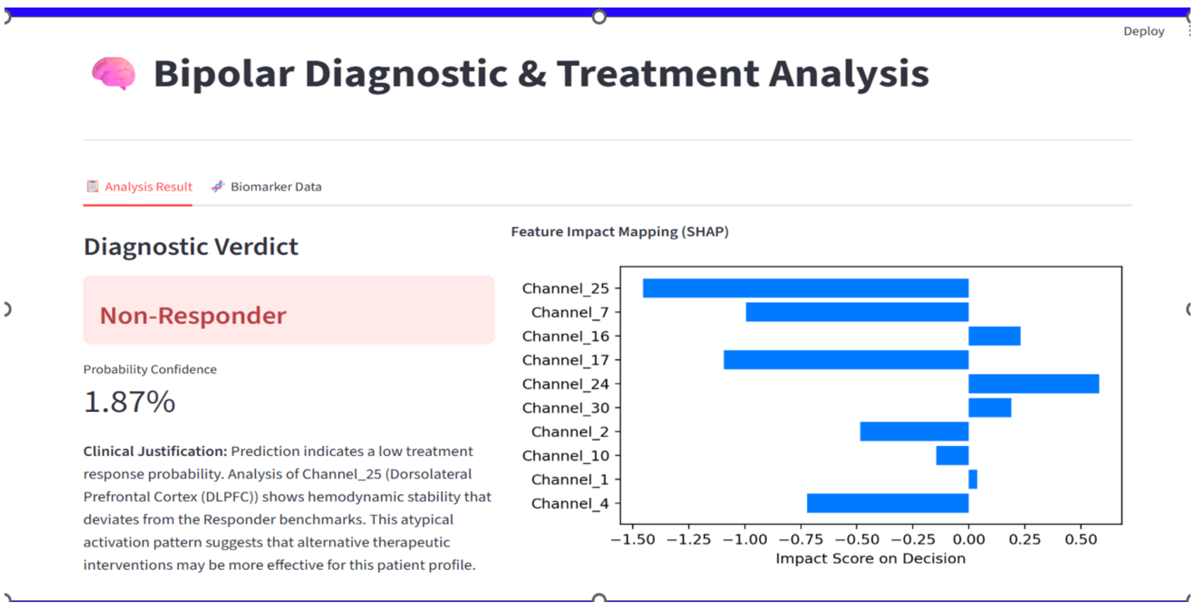
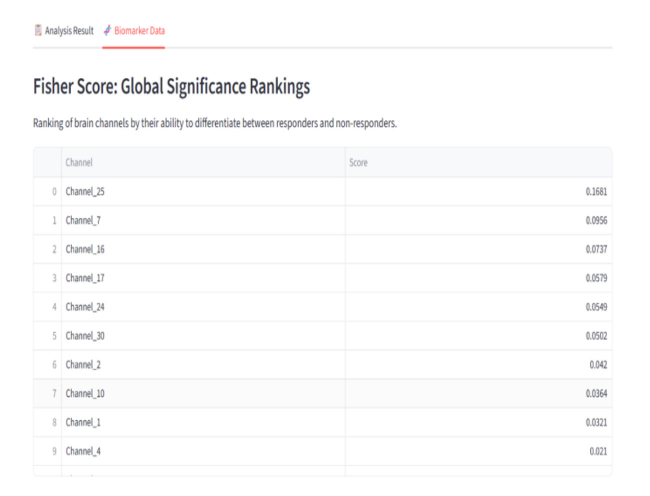


Fig.8 RESULT

The present system relies on its successes to demonstrate the limitations of the current system. Due to the current systems reliance on both the quantity and quality of data that it can be fed into the model. Of course, if a large enough amount of high-quality data could be come into the system, the model would be able to provide the highest level potential of performance possible. Therefore, the current system has limitations due to its restriction to a small maximum amount of "advanced" data being added to the current models set of "data" features. The addition of some type of testing using larger data sets and integrating the additional advanced algorithms with the current configuration of the current system could allow for improvements to the robustness and relyability of the current system. The removal of restrictions on the current system will further support the already proven success of the current model.


Analysis Result Biomarker Data

Fisher Score: Global Significance Rankings

Ranking of brain channels by their ability to differentiate between responders and non-responders.

	Channel	Score
0	Channel_25	0.1681
1	Channel_7	0.0956
2	Channel_16	0.0737
3	Channel_17	0.0579
4	Channel_24	0.0549
5	Channel_30	0.0502
6	Channel_2	0.042
7	Channel_10	0.0364
8	Channel_1	0.0321
9	Channel_4	0.021

Fig.9 FISHER SCORE SELECTION

To summarize, the proposed model successfully illustrates how machine learning may be utilized by a clinician to forecast how he/she will respond to a treatment for bipolar disorder. The combination of fNIRS data (fNIRS is a functional near infrared spectroscopic system) along with the use of feature selection and the XGBoost method yields accurate and efficient results in this research. The prototype contains an easy to use interface that clinicians can readily utilize in their day-to-day activities practice as well as a prototype that can be employed in real-world settings.

The data analyzed in this research emphasizes the need for a more data-driven method to medical care. This project has demonstrated how data will assist in developing better personalized treatments and making appropriate clinical choices. Overall, this has demonstrated to possess great potential for future research and clinical use alike.

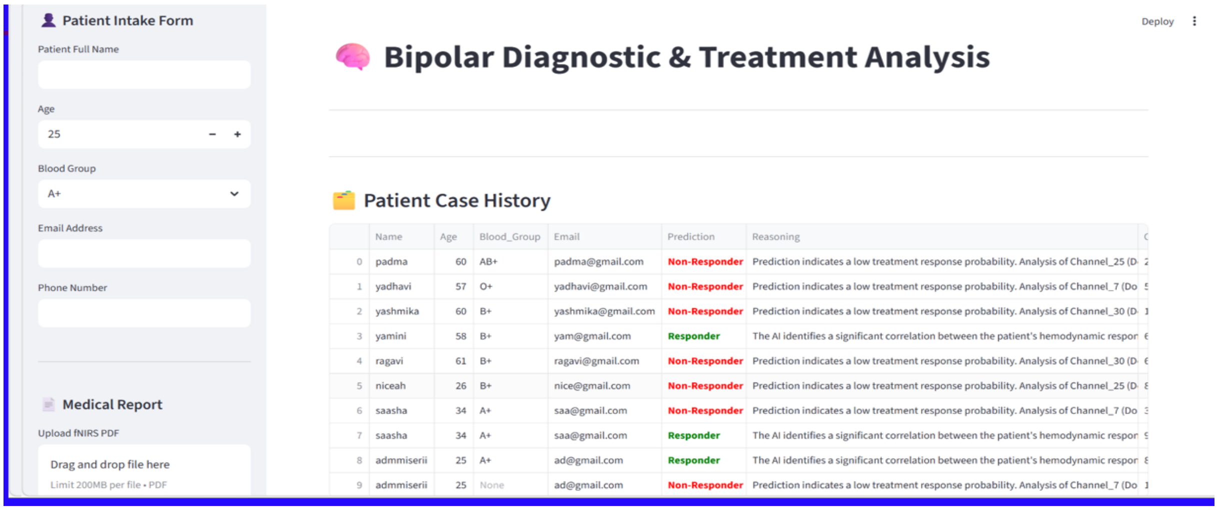


Fig.10BIPOLAR DIAGNOSTIC PATIENT HISTORY

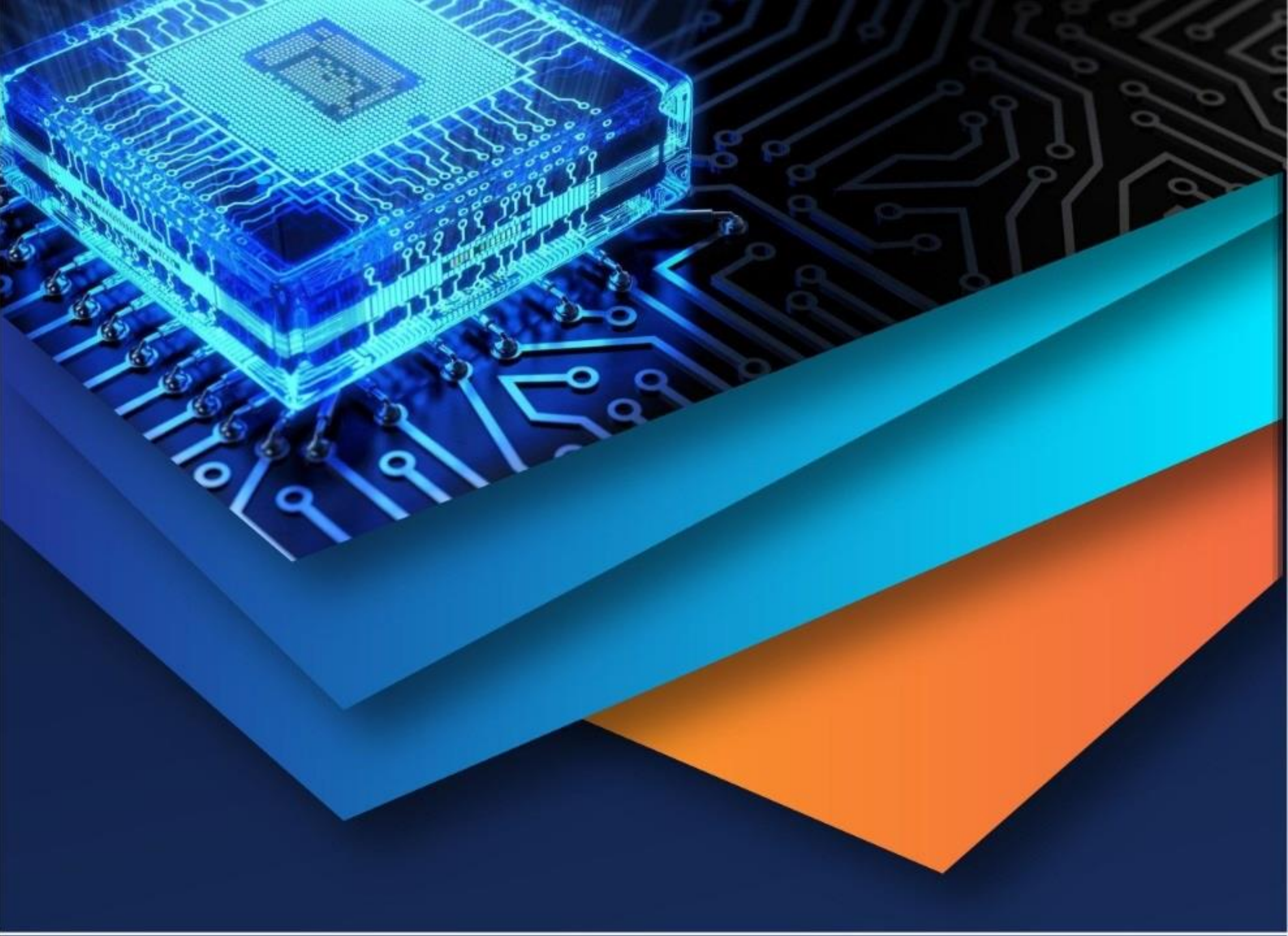
VIII. CONCLUSION

The purpose of this study is to demonstrate the application of machine learning as a method of assisting with the treatment of individuals suffering from BPD via various methods: fNIRS brain signal data, pre-processing, extracting features from that data, and finally selecting important features for predicting response to therapy using the Fisher Score. Then, using the XGBoost machine learning algorithm to make predictions about whether or not an individual who has BPD is going to respond to tPBM therapy. The entire system has been developed to be able to use a web interface so that physicians can upload data to the system and obtain predictions without any reservations. Finally, this project will enable more effective and efficient clinical decisions to be made.

Furthermore, this study presents evidence of how technology aids in providing better healthcare services through the provision of individualized care. It allows for predicting the treatment outcome before administering it, rather than just using a trial-and-error process, which saves time and avoids providing unnecessary treatment. The visualization and storage of the data enhances its utility and ease of understanding. Technology will be improved upon with continued data accumulation and development of improved algorithms; however, as seen in this project, there are already many indications that an improvement in performance will occur over time. The future of the approach will include being adapted into real medical environments. Ultimately, this project has positively impacted the effectiveness and quality of care provided to people with bipolar disorder.

REFERENCES

[1] John D. Kelleher, Fundamentals of Machine Learning for Predictive Data Analytics, MIT Press, 2015.
[2] Tianqi Chen and Carlos Guestrin, "XGBoost: A Scalable Tree Boosting System," Proceedings of the ACM SIGKDD, 2016.
[3] Artur M. R. Gomes et al., "Functional Near-Infrared Spectroscopy in Mental Health Research," NeuroImage, 2019.
[4] World Health Organization, "Bipolar Disorder," 2023.
[5] George B. Moody, "Biomedical Signal Processing and Analysis," IEEE Press, 2017.
[6] Isabelle Guyon, "An Introduction to Variable and Feature Selection," Journal of Machine Learning Research, 2003.
[7] Leo Breiman, "Statistical Modeling: The Two Cultures," Statistical Science, 2001.
[8] Hinton Geoffrey, "Reducing the Dimensionality of Data with Neural Networks," Science, 2006.
[9] Pinti Paola et al., "The Present and Future Use of Functional Near-Infrared Spectroscopy," Neurophotonics, 2018.
[10] IEEE, "Machine Learning in Healthcare Applications," IEEE Journals, 2020.



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY