# ANALYSIS AND CLASSIFICATION OF DISORDERS OF CONSCIOUSNESS USING MACHINE LEARNING METHODS BASED ON SCALP EEG

***SREELAKSHMI RAVEENDRAN***
***2019PHD12003***




Indian Institute of Information Technology Kottayam

# ANALYSIS AND CLASSIFICATION OF DISORDERS OF CONSCIOUSNESS USING MACHINE LEARNING METHODS BASED ON SCALP EEG

**Thesis**

Submitted in Partial Fulfillment of the

Requirements for the Degree of

**Doctor of Philosophy**

*by*

***SREELAKSHMI RAVEENDRAN***

2019PHD12003




DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

**INDIAN INSTITUTE OF INFORMATION TECHNOLOGY KOTTAYAM**

PALA - 686635, KERALA, INDIA

April 2025

भारतीय सूचना प्रौद्योगिकी संस्थान कोट्टायम
INDIAN INSTITUTE OF INFORMATION TECHNOLOGY KOTTAYAM
(An Institution under Ministry of Education, Govt. of India governed by an Act of Parliament)
Tel: 04822-202 100 (office), 202135 (Academics), 202110 (Security), 202102 (Accounts)
Email: registrar@iitkotayam.ac.in, Website: www.iitkottayam.ac.in

# CERTIFICATE

**Date:** 25 / 04 / 2025

This is to certify that the thesis entitled **Analysis and Classification of Disorders of Consciousness using Machine Learning Methods based on Scalp EEG** submitted by **Sreelakshmi Raveendran (2019PHD12003)** to the **Indian Institute of Information Technology Kottayam** is a record of bonafide research work carried out by him under our supervision and is worthy of consideration for the award of the degree of **Doctor of Philosophy** of the Institute. The results contained in this Thesis have not been submitted in part or full to any other university or institute for any Degree or Diploma.

| **Guide** | **Co-Guide** | **Co-Guide** |
|---|---|---|
| **Dr. Kala S** | **Dr. Jayakrushna Sahoo** | **Dr. Santhos Kumar** |
| *Assistant Professor* | *Assistant Professor* | *Assistant Professor* |
| Department of ECE | Department of CSE | Department of ECE |
| IIIT Kottayam | IIIT Kottayam | IIIT Kottayam |

# DECLARATION

I declare that

(a) the work contained in the thesis is original and has been done by me under the guidance of my supervisors;

(b) the work has not been submitted to any other institute for any degree or diploma;

(c) I have followed the guidelines provided by the Institute in preparing the thesis;

(d) I have conformed to ethical norms and guidelines while writing the thesis;

(e) whenever I have used materials (data, models, figures, and text) from other sources, I have given due credit to them by citing them in the text of the thesis, giving their details in the references, and taking permission from the copyright owners of the sources, if necessary, and;

(f) whenever I have quoted written materials from other sources, I have put them under quotation marks and given due credit to the sources by citing them and providing required details in the references.

(g) The thesis has been checked for plagiarism with the tool recommended by the institute and has been found to confirm the plagiarism policy of the institute.



**Place:** IIIT Kottayam

**Date:** 25 / 04 / 2025

**Sreelakshmi Raveendran**

**2019PHD12003**

**Department of ECE**

© Copyright 2025 by Indian Institute of Information Technology Kottayam
All rights reserved

# Acknowledgments

I am deeply indebted to numerous individuals whose support and encouragement have been instrumental throughout my academic journey at IIIT Kottayam. The successful completion of this doctoral program would not have been possible without the assistance and cooperation of many. I would like to express my sincere gratitude to all those who contributed to my progress and growth over the course of my research.

First and foremost among them is my supervisor, Dr. Kala S, whose invaluable guidance, encouragement, and unwavering support have been instrumental throughout the course of this research. I am profoundly grateful for the opportunity to pursue my research under her guidance in the Department of ECE. Her approachability and positive demeanor fostered a collaborative and intellectually stimulating research environment, making my time in Intelligent Integrated Circuits and Systems ($I^2CS$) Research Lab, both productive and enriching. Her research intuition and unwavering commitment to excellence greatly inspired me. Her insightful feedback and high standards for technical writing significantly shaped the quality and productivity of my PhD work.

I would also like to express my sincere gratitude to my co-guides,
Dr. Jayakrushna Sahoo, Assistant Professor in the Department of CSE, and Dr. Santhos Kumar, Assistant Professor in the Department of ECE. Their constructive suggestions, technical expertise, and continuous support played a crucial role in strengthening the interdisciplinary aspects of my research.

I would like to sincerely thank my Doctoral Committee (DC) members, Dr. P. Victer Paul, Assistant Professor, Department of CSE, and Dr. J. V. Bibal Benifa, Assistant Professor, Department of CSE for their encouragement and constructive input at various stages of my research work.

I would like to thank Prof. Prasad Krishna, Director, IIIT Kottayam, Dr. M Radhakrishnan, Registrar, IIIT Kottayam, Prof P. Mohanan, Professor In-charge (Academics), IIIT Kottayam, and Dr. Ananth A, Head, Department of ECE, IIIT Kottayam, for their administrative support to carry out research work. I also thank all the faculty members in the Department of ECE for their co-operation during my PhD program. Heartfelt thanks to my wonderful lab mates Sweety, Prasha, Remya, Aparna and Asif — without their presence, support, and companionship, my PhD journey would have been incomplete. Their support, laughter, and friendship

brought light to even the toughest days.

I am deeply grateful to Prof. A. G. Ramakrishnan, Adjunct Faculty, Department of Heritage Science and Technology, IIT Hyderabad for his unwavering support and trust throughout my research journey. His consistent encouragement and valuable guidance have been instrumental in the successful completion of this thesis. His insights and assistance, particularly during the preparation of research papers were invaluable. The numerous discussions with him greatly contributed to shaping the direction of my research. I remain sincerely indebted to him for his exceptional mentorship and support.

I would like to extend my heartfelt gratitude to the team from the National Institute of Mental Health and Neurosciences (NIMHANS), Bangalore, coordinated by Dr. Subasree R, Additional Professor, Department of Neurology, NIMHANS, Bangalore for their invaluable support in the data collection and clinical evaluation phases of this research. The in-depth formulation of this study, particularly with the specific dataset employed, would not have been possible without her guidance and initiative. I am also deeply thankful to Dr. Raghavendra Kenchaiah, Department of Neurology, NIMHANS, Bangalore, Dr. Farsana M. K, Department of Neurology, NIMHANS, Bangalore, Dr. Ravindranath Choudary Mudlamuri, Department of Neurology, NIMHANS, Bangalore, Dr. Sonia Bansal, Department of Neuroanaesthesia and Neurocritical Care, NIMHANS, Bangalore, and Dr. Binu V. S, Department of Biostatistics, NIMHANS, Bangalore for their dedicated contributions and clinical expertise. Their support was instrumental in ensuring the scientific rigor and clinical relevance of this work.

I would like to express my sincere gratitude to my parents Raveendranath C. R and Sarala K.M, my brother Srinath Raveendran, and my sister-in-law Dr. Saranya for their persistent support during my Ph.D. program. Finally, I cannot forget thanking the Almighty for making me stronger and perseverant in difficult situations.

This thesis is dedicated to my late grandmother Mrs. Bharathi whose medical condition became the sole reason for this research.

**Sreelakshmi Raveendran**

*Dedicated to my parents for their endelss support and my grandmother Bharathi whose medical condition became the reason for this research...*

# Abstract

Disorders of consciousness (DOC), encompassing conditions namely coma, unresponsive wakefulness syndrome, and minimally conscious state, pose significant challenges in clinical neuroscience. Behavioral assessments, though widely employed in practice for the categorization of DOC cohorts, are inherently subjective and susceptible to misinterpretation. These limitations contribute to a misdiagnosis rate of approximately 37– 40%. The integration of neuroimaging techniques, such as electroencephalography (EEG), has the potential to significantly reduce this rate by providing objective measures of brain function. Resting-state EEG (rs-EEG) has emerged as a promising non-invasive modality to capture neural activity associated with consciousness. However, EEG signals are inherently high-dimensional, noisy, and complex, often requiring advanced computational techniques for meaningful interpretation. Further challenges arise from inter-patient variability, heterogeneity in injury etiology, and confounding factors such as medication effects and fluctuating arousal levels. The lack of standardized preprocessing protocols and the opacity of many machine learning models limit their adoption in clinical settings. This research aims to address these challenges by developing a robust, interpretable, and clinically relevant machine learning framework for assessing and predicting consciousness levels in DOC patients using rs-EEG data. A high-quality EEG dataset was acquired in collaboration with the National Institute of Mental Health and Neurosciences, Bangalore, including patients with DOC of diverse etiologies as well as healthy control subjects. This dataset serves as the foundation for the development, validation, and benchmarking of novel computational models implemented during the research.

In the initial phase of the work, variational mode decomposition (VMD) is employed as a preprocessing technique to improve the classification ability of EEG-derived features. VMD components enable the extraction of nuanced features that better reflect the underlying neural dynamics. A broad range of features derived from VMD components is used to train and evaluate various machine learning classifiers. Comparative analysis against conventional frequency band features and raw EEG signals demonstrates the diagnostic ability of VMD-based preprocessing in classifying DOC subtypes.

To assess the consciousness level of each subject based on EEG-derived features, this research proposes a novel ensemble of ensemble regressor model, by combining

the strengths of extreme gradient boosting, random forest, and support vector regression within a stacking ensemble framework. To enhance interpretability, shapley additive explanations (SHAP) were utilized to identify influential features across the ensemble, which were then used to construct a simplified linear model for estimating a normalized consciousness score aligned with the coma recovery scale–revised.

To further understand the underlying neural activity of DOC cohorts, functional connectivity (FC) analysis using metrics such as phase lag index (PLI), weighted phase lag index, debiased weighted phase lag index (dwPLI) amplitude envelope correlation (AEC), and coherence were performed. Based on these connectivity features, a series of deep learning models including multilayer perceptrons, convolutional neural networks (CNN), recurrent neural networks, gated recurrent units, long short-term memory (LSTM) networks, and a hybrid CNN-LSTM were implemented to classify DOC cohorts. AEC metric demonstrated better performance in discriminating DOC subgroups. Further, graph theoretical analysis was performed to assess the topological organization of brain networks. Undirected networks were generated using selected FC measures namely dwPLI, coherence, and AEC. Network properties such as modularity, small-worldness, and betweenness centrality were examined for their associations with clinical outcomes, including the glasgow outcome scale–extended (GOSE). Together, these findings contribute to the development of interpretable, data-driven tools for improving diagnostic accuracy and enhancing clinical decision-making in DOC assessment.

**Keywords:** Disorders of consciousness, variational mode decomposition, SHAP, ensemble models, functional connectivity, graph theory, canonical correlation analysis, sliding window correlation.

# Contents

**Abstract** **xv**

**List of Figures** **xxi**

**List of Tables** **xxiii**

**Nomenclature** **xxvii**

**Acronyms** **xxxii**

**1 INTRODUCTION** **1**

1.1 Motivation, Objectives, and Scope . . . . . 3
1.2 Behavioural Assessment Methods . . . . . 4
1.3 Resting State EEG Paradigms for DOC . . . . . 5
1.4 Automated Classification of DOC Patients . . . . . 7
1.5 Contribution of Thesis . . . . . 8
1.6 Organization of Thesis . . . . . 9

**2 Literature Review** **12**

2.1 Resting-state EEG Signal Paradigms for Diagnosis of DOC Patients 13
2.2 Spectral and Nonlinear Analysis of EEG . . . . . 13
2.3 Functional Connectivity and Graph Theoretical Metrics . . . . . 14
2.4 Graph Theory and Network Topology . . . . . 15
2.5 Advanced Signal Decomposition Techniques . . . . . 16
2.6 Machine Learning and Deep Learning Applications . . . . . 17

**3 EEG Signal Acquisition and Standardization** **19**

3.1 Experimental Setup . . . . . 19
3.2 Inclusion/Exclusion Criteria . . . . . 20
3.3 Details of the EEG System Used . . . . . 21
3.4 Demographics of the Subjects . . . . . 22

3.5 Quantitative and Qualitative Analysis of Data . . . . . 22
3.6 Ethical Clearance . . . . . 22
3.7 Sample Size Calculation . . . . . 23

**4 Variational Mode Decomposition based EEG Signal Preprocessing and Classification by Machine Learning 26**
4.1 Feature Extraction . . . . . 27
4.1.1 Spectral Power . . . . . 27
4.1.2 Nonlinear dynamic features . . . . . 29
4.1.3 Spectral Entropy . . . . . 32
4.1.4 Kurtosis . . . . . 33
4.1.5 Skewness . . . . . 34
4.2 Variational Mode Decomposition . . . . . 35
4.3 Methodology . . . . . 37
4.4 Classifier Models . . . . . 39
4.5 Results . . . . . 40
4.5.1 Approach 1: Features Derived from the Raw EEG . . . . . 40
4.5.2 Approach 2: Features Derived from the Frequency Bands . . 43
4.5.3 Approach 3: Features Derived from the VMD Modes . . . . 45
4.5.4 Performance Comparison Between the Three Approaches . . 47
4.6 Statistical Analysis . . . . . 48
4.6.1 Clinical Implications of Correlations . . . . . 49
4.6.2 Improvement in the Performance of the Classifiers Due to Feature Selection . . . . . 54
4.6.3 Performance Comparison Between the Three Approaches Post Feature Selection . . . . . 55
4.7 Summary of the Chapter . . . . . 61

**5 Regression Models to Predict Consciousness Level 63**
5.1 Feature Extraction . . . . . 64
5.2 Prediction Methods and Performance Metrics . . . . . 66
5.2.1 Prediction Methods . . . . . 66
5.2.2 Performance Metrics . . . . . 67
5.3 Ensemble of Ensemble Regressor Model: EoERM . . . . . 68
5.4 Explainability of Features and Feature Selection using SHAP . . . . 70
5.5 Consciousness Level Estimation . . . . . 71
5.6 Results . . . . . 74

5.6.1 Performance of Basic Regressor Models to Predict Consciousness Level . . . 74
5.6.2 Results of the Proposed Model without Feature Selection . . 76
5.6.3 Results of the Proposed Model with Feature Selection . . . . 76
5.6.4 Normalized Consciousness Level-Based Analysis . . . . . 78
5.7 Summary of the Chapter . . . 81

**6 Functional Connectivity Analysis and Deep Learning-Based Classification of DOC Patients 82**
6.1 Feature Extraction . . . 83
6.1.1 Functional Connectivity Measures . . . 84
6.1.2 Graph Theory Metrics . . . 88
6.1.3 Sliding Window Correlation . . . 90
6.1.4 Canonical Correlation Analysis . . . 91
6.2 Choice of ANN Models for Multiclass Classification . . . 91
6.2.1 MLP Models . . . 92
6.2.2 Sequential Models: Simple RNN, LSTM, and GRU . . . 93
6.3 Methodology . . . 95
6.3.1 Functional Connectivity based Classification using Deep Learning Models . . . 95
6.3.2 Graph Theoretical Analysis of Undirected Weighted Brain Networks . . . 97
6.4 Results . . . 97
6.4.1 Performance Evaluation of Deep Learning Models . . . 97
6.4.2 Graph Theoretical Network Analysis of Consciousness States 120
6.5 Comparison with Existing Work . . . 124
6.6 Summary of the Chapter . . . 127

**7 Conclusion and Future Work 128**
7.1 Conclusions . . . 128
7.2 Future Directions . . . 129

**List of Publications 130**

**References 133**

**Curriculum Vitae 147**

# List of Figures


3.1 Resting-state EEG acquisition system . . . . . 20
3.2 Resting-state EEG recording setup . . . . . 21

4.1 Block diagram of proposed approach. . . . . 38
4.2 Mean sample entropy values . . . . . 41
4.3 Mean spectral entropy values . . . . . 42
4.4 Mean kurtosis values . . . . . 42
4.5 Mean skewness values . . . . . 43
4.6 Spearman rank correlation coefficients between the CRS-R scores and selected features extracted from each of the 17 EEG channels . 50
4.7 Spearman rank correlation coefficients between the GCS scores and selected features extracted from each of the 17 EEG channels . . . . 50
4.8 Spearman rank correlation coefficients between the CRS-R scores and selected features extracted from each frequency bands . . . . . 51
4.9 Spearman rank correlation coefficients between the GCS scores and selected features extracted from each frequency bands . . . . . 51
4.10 Spearman rank correlation coefficients between the CRS-R scores and selected features extracted from each of the five VMD modes . 52
4.11 Spearman rank correlation coefficients (y-axis) between the GCS scores and selected features extracted from each of the five VMD modes . . . . . 53

5.1 Graphical depiction of the proposed methodology . . . . . 70
5.2 Summary plot obtained using SHAP method for random forest regressor model . . . . . 78
5.3 Summary plot obtained using SHAP method for SVR regressor model 78
5.4 Summary plot obtained using SHAP method for XGBoost regressor model . . . . . 79
5.5 Normalised consciousness levels for the three groups of DOC patients. 80

6.1 Proposed methodology: FC-based multiclass classification of DOC patients using various deep learning models . . . . . 96
6.2 Variations of graph metrics from AEC network between groups . . . 121
6.3 Variations of graph metrics from dwPLI network between groups . . 122
6.4 Variations of graph metrics from coherence network between groups 123
6.5 Canonical weights for graph theory features from dwPLI network . 125

# List of Tables


3.1 Demographics of the UWS group at baseline . . . . . 23
3.2 Demographics of the coma group at baseline . . . . . 24
3.3 Demographics of the MCS group at baseline . . . . . 25

4.1 Binary (UWS vs MCS) classification results (in %) using 10-fold cross-validation with all four features directly extracted from the 17 channels of the raw EEG. Number of patients: 30 . . . . . 42
4.2 Multi-class (Coma vs UWS vs MCS) classification results (in %) using 10-fold cross-validation with all the four features extracted from the 17 channels of the raw EEG. Number of patients: 45 . . . 43
4.3 Binary (UWS vs MCS) classification results (in %) using 10-fold cross-validation with the entire set of frequency band-based features. Number of patients: 30. The performance is inadequate with all the classifiers . . . . . 45
4.4 Multi-class (Coma vs UWS vs MCS) classification results (in %) using 10-fold cross-validation for the entire set of frequency band-derived features. Number of patients: 45. The performance is unsatisfactory with all the classifiers . . . . . 45
4.5 Binary (UWS vs MCS) classification results (in %) using 10-fold cross-validation employing all the VMD-based features. Number of patients: 30. Ensemble bagged tree performs the best with VMD-based features . . . . . 46
4.6 Multi-class (Coma vs UWS vs MCS) classification results (in %) using 10-fold cross-validation employing all the VMD-based features. Number of patients: 45. Ensemble bagged tree performs the best with VMD-based features . . . . . 47
4.7 Comparison of percentage increase in accuracy of VMD-derived features over frequency band-derived (FB) and raw EEG-derived (EEG) features, when all the features are employed in each of the approaches . . . . . 48

4.8 Binary (UWS vs MCS) classification results (in %) using 10-fold cross-validation with significant features extracted using Kruskal-Wallis test for the 3 different approaches. Number of patients: 30 . 55
4.9 Multi-class (Coma vs UWS vs MCS) classification results (in %) using 10-fold cross-validation with significant features extracted using Kruskal-Wallis test for the 3 different approaches. Number of patients: 45 . . . . . 56
4.10 Improvement in accuracy (%) of different classifiers due to the selection of features shown to be statistically significant by Kruskal-Wallis test: VMD-derived features. (IA: absolute improvement in accuracy) . . . . . 56
4.11 Comparison of percentage increase in accuracy of VMD over frequency band-derived (FB) and raw EEG-derived (EEG) features after feature selection based on Kruskal Wallis test . . . . . 57
4.12 Comparison of performance of proposed method with existing work 59

5.1 Combined feature importance from each of the 3 regressors calculated using SHAP . . . . . 72
5.2 Comparison of different regression models based on performance metrics . . . . . 74
5.3 Performance of the proposed regressor EoERM with and without feature selection task, as well as normalized consciousness level calculated using the custom equation. . . . . 81

6.1 Performance comparison of different ANN models for multiclass classification with 20 second window length and 80/20 train-test split (coma vs. UWS vs. MCS vs. HC) using three different features and a combination of two of them. . . . . 98
6.2 Accuracy with PLI features for different window sizes with 80-20 fixed train-test split . . . . . 99
6.3 Best fold obtained using group k-fold cross validation and the performance metrics of each model using PLI features . . . . . 101
6.4 Accuracy with wPLI features for different window sizes with 80-20 fixed train-test split . . . . . 104
6.5 Best fold obtained using group k-fold cross validation and the performance metrics of each model using wPLI features . . . . . 105
6.6 Accuracy with AEC features for different window sizes with 80-20 fixed train-test split . . . . . 108

6.7 Best fold obtained using group k-fold cross validation and the performance metrics for each model using AEC features. . . . . 109
6.8 Accuracy with AEC+wPLI features for different window sizes with 80-20 train test split . . . . . 111
6.9 Best performing fold in group k-fold cross validation and the performance metrics for each model using combined AEC and wPLI features. . . . . 113
6.10 Confidence interval for the implemented models using different features 115
6.11 Confidence interval for the implemented models using different features 116
6.12 Comparison of our performance with the techniques in the literature. (TBI: traumatic brain injury) . . . . . 126

# Nomenclature

| | |
|---|---|
| $\Delta\phi(t)$ | Instantaneous phase difference |
| $\delta(t) + j/\pi t$ | Hilbert transform kernel |
| $\epsilon_a$ | Absolute tolerance |
| $\epsilon_r$ | Relative tolerance |
| $\hat{s}(t)$ | Discrete symbol |
| $\hat{u}_k(\omega)$ | Oscillatory mode in frequency domain |
| $\hat{x}(t)$ | Reconstructed signal |
| $\hat{y}_i$ | Predicted value |
| $\lambda(t)$ | Lagrange multiplier |
| $\lambda, \alpha$ | Regularization parameter, penalty factor |
| $\mu$ | Mean |
| $\omega_k$ | Center frequency |
| $\partial t$ | Partial derivative |
| $\phi_x(t)$ | Instantaneous phase |
| $\pi_i$ | Ordinal pattern |
| $\sigma V^2$ | Microvolts squared |
| $\sigma$ | Standard deviation |
| $\tau$ | Embedding time delay |
| $A^m(r)$ | Probability that two sequences are similar for $m+1$ points |
| $A_f(t)$ | Amplitude envelope of filtered signal |

| | |
|---|---|
| $A_x(t)$ | Instantaneous amplitude |
| $ab$ | Absolute power |
| $Acc.$ | Accuracy |
| $B^m(r)$ | Probability that two sequences are similar for $m$ points |
| $C_i$ | Clustering coefficient of $i^{th}$ node |
| $C_{rand}$ | Clustering coefficient of randomly generated network |
| $C_{rs}(f)$ | Coherence between $r(t)$ and $s(t)$ |
| $D$ | Embedding dimension |
| $d_{ij}$ | Path length between two nodes |
| $E_S$ | Spectral Entropy |
| $f(t)$ | Continuous time EEG signal |
| $F1-S$ | F1-Score |
| $g$ | Data points |
| $H(x(t))$ | Hilbert transform |
| $J$ | Cost function |
| $k$ | Number of oscillatory modes |
| $L$ | Path length |
| $L_C$ | Consciousness level |
| $L_{rand}$ | Path length of randomly generated network |
| $m$ | Run length |
| $m+1$ | Next point |
| $N$ | Number of time samples |
| $P(f)$ | Probability distribution |
| $P_i$ | Probability of occurence of $i^{th}$ symbol |
| $P_{rr}(f)$, $P_{rr}(f)$ | Spectral density of individual signals |

| | |
|---|---|
| $P_{rs}(f)$ | Cross spectral density |
| $Pr.$ | Precision |
| $q$ | q-dimension space |
| $r$ | Tolerance window |
| $ReC.$ | Recall |
| $rp$ | Relative power |
| $S$ | Small-worldness |
| $S_X$ | Symbolic sequence of target |
| $S_Y$ | Symbolic sequence of source |
| $S_k$ | Skewness |
| $u_k(t)$ | Oscillatory modes |
| $X$ | Source process |
| $x(n)$ | Discrete time EEG signal |
| $x_t^{(k)}, y_t^{(k)}$ | State vectors |
| $Y$ | Target process |
| $z(t)$ | Hilbert transform |
| $\mathrm{Im}(\cdot)$ | Imaginary component |
| K | Kurtosis |
| $\mathrm{R}^2$ | Coefficient of Determination |
| $\mathrm{Re}(\cdot)$ | Real component |
| SampEn | Sample Entropy |

# Acronyms

**ADMM** Alternating Direction Method of Multipliers

**AEC** Amplitude Envelope Correlation

**AI** Artificial Intelligence

**ApEn** Approximate Entropy

**BPL** Below Poverty Line

**CCA** Canonical Correlation Analysis

**CEEMD** Complementary Ensemble Empirical Mode Decomposition

**CI** Confidence Interval

**CNN** Convolutional Neural Networks

**CNN-LSTM** Convolutional Neural Networks-Long Short-Term Memory

**CRS-R** Coma Recovery Scale - Revised

**CT** Computed Tomography

**CV** Cross Validation

**dFC** Dynamic Functional Connectivity

**DL** Deep Learning

**DMN** Default Mode Network

**DOC** Disorders of Consciousness

**DT** Decision Tree

**DTF** Directed Transfer Function

**dwPLI** debaised weighted Phase Lag Index

**EBT** Ensemble Bagged Tree

**EEG** Electroencephalogram

**EMD** Empirical Mode Decomposition

**EoERM** Ensemble of Ensemble Regressor Model

**ERPs** Event Related Potentials

**FB** Frequency Band

**FC** Functional Connectivity

**FIR** Finite Impulse Response

**fMRI** Functional Magnetic Resonance Imaging

**GCS** Glasgow Coma Scale

**GOSE** lasgow Outcome Scale–Extended

**GRU** Gated Recurrent Units

**HC** Healthy Controls

**IC** Imaginery Coherence

**IMF** Intrinsic Mode Function

**k** Number of Intrinsic Modes

**KCC** K-Complex Consistency

**KNN** k-Nearest Neighbor

**LDA** Linear Discriminant Analysis

**Li.SVM** Linear Support Vector Machine

**LORETA** Low-Resolution Brain Electromagnetic Tomography

**LSTM** Long Short-Term Memory

**LZC** Lempel-Ziv Complexity

**MAE** Mean Absolute Error

**MCS** Minimally Conscious State

**MCS+** Minimally Conscious State+

**MCS−** Minimally Conscious State−

**MEG** Magnetoencephalography

**MI** Mutual Information

**ML** Machine Learning

**MLP** Multilayer Perceptrons

**MRI** Magnetic Resonance Imaging

**MSE** Mean Squared Error

**NIMHANS** National Institute of Mental Health and Neurosciences

**PCA** Principal Component Analysis

**PCoh** Partial Coherence

**PE** Permutation Entropy

**PLI** Phase Lag Index

**PLV** Phase-Locking Value

**PSD** Power Spectral Density

**Rb.SVM** Radial Basis Support Vector Machine

**ReLU** Rectified Linear Unit

**RFE** Recursive Feature Elimination

**RNN** Recurrent Neural Networks

**rs-EEG** Resting-State Electroencephalogram

**SCS** Spinal Cord Stimulation

**SHAP** Shapley Additive Explanations

**SMART** Sensory Modality Assessment and Rehabilitation Technique

**std.SWC** Standard Deviation of SWC

**STE** Symbolic Transfer Entropy

**SVM** Support Vector Machine

**SVR** Support Vector Regression

**SWC** Sliding Window Correlation

**TBI** Traumatic Brain Injury

**tDCS** Transcranial Direct Current Stimulation

**TE** Transfer Entropy

**UWS** Unresponsive Wakefulness Syndrome

**VMD** Variational Mode Decomposition

**wPLI** weighted Phase Lag Index

**wSMI** weighted Symbolic Mutual Information

**XAI** Explainable Artificial Intelligence

**XGBoost** Extreme Gradient Boosting Model

# Chapter 1

# INTRODUCTION

---

Disorders of consciousness (DOC) comprise a spectrum of clinical conditions characterized by significant impairments in arousal and awareness [1]. These conditions present complex diagnostic and therapeutic challenges for patients, their families, and healthcare providers [2]. A comprehensive understanding of the various DOC categories and their distinguishing features is essential for accurate diagnosis, appropriate treatment planning, and prognostic evaluation [3, 4]. One of the most recognized forms of DOC is coma, which represents a state of profound unconsciousness. Individuals in a comatose state are unresponsive to external stimuli, exhibit no signs of awareness, and typically have closed eyes, with no evidence of wakefulness or purposeful behavior [5]. Coma is often precipitated by severe traumatic brain injury, cerebrovascular accidents, intracranial hemorrhages, or other pathological conditions that significantly impair normal cerebral function [6]. Another DOC category is unresponsive wakefulness syndrome (UWS) [7]. Patients with UWS may display sleep-wake cycles, open their eyes, and maintain basic autonomic and physiological functions [8, 9]. However, despite these behaviors, there is a complete absence of conscious awareness and purposeful interaction with the environment [9]. The third classification within the DOC continuum is the minimally conscious state (MCS) [9]. Individuals in MCS demonstrate

limited, but reproducible, signs of awareness and wakefulness [10]. Unlike UWS, patients in MCS may exhibit intermittent purposeful behaviors such as following simple commands, gesturing, verbalization, or displaying emotional expressions in response to external stimuli [11]. However, the presence of awareness in MCS is often inconsistent and subject to fluctuation, posing significant challenges for reliable clinical assessment [10].

Categorizing patients along the DOC spectrum typically relies on standardized behavioral assessment tools [12]. Among these, the coma recovery scale–revised (CRS-R) is widely recognized as the most reliable instrument for evaluating the level of consciousness and tracking recovery trajectories in DOC patients [13]. This method evaluate observable responses like eye movements, vocalizations, and motor commands, providing valuable clinical insights. However, behavioral assessments have inherent limitations. They are subjective, influenced by examiner expertise, patient arousal fluctuations, and the subtlety of responses [14]. These methods often fail to detect covert consciousness a condition where patients exhibit neural markers of awareness despite the absence of overt behavioral signs. Research indicates that up to 37– 40% of patients diagnosed with an unresponsive state may possess covert consciousness, underscoring the risk of misdiagnosis and its implications for treatment and prognosis [14, 15]. The shortcomings of behavioral evaluations highlight the need for objective, neurophysiological measures to complement clinical assessments. Resting-state electroencephalography (EEG) offers a promising avenue, capturing subtle neural activity patterns that may elude behavioral observation [16]. However, the complexity and the amount of EEG data necessitate advanced computational methods for effective analysis. In this context, machine learning and deep learning models have emerged as transformative tools, enabling the identification and classification of EEG patterns with better precision [17].

## 1.1 Motivation, Objectives, and Scope

The misdiagnosis error arising due to behavioral assessment method, limited availability of comprehensive EEG datasets specific to DOC cohorts, as well as the scarcity of robust, EEG feature-driven classification models and predictive frameworks for assessing levels of consciousness are challenges exisitng in this research field. This research addresses these limitations by developing advanced, explainable machine learning techniques to improve the classification and consciousness assessment of DOC patients using resting-state EEG data.

The objective of this research is to advance the understanding, diagnosis, and assessment of consciousness levels in DOC patients through a comprehensive analysis of resting-state EEG signals. The study aims to investigate and classify neural patterns associated with varying states of consciousness using established and emerging computational methodologies. For that, a high-quality dataset comprising resting-state EEG recordings from DOC patients with diverse etiologies, along with signals from healthy controls, has been created in collaboration with the National Institute of Mental Health and Neurosciences (NIMHANS), Bangalore. This dataset forms the foundation for this research.

As a first step, this research investigates the use of variational mode decomposition (VMD) as a signal preprocessing technique to improve the performance of the implemeted classifier models. To validate the utility of VMD-based preprocessing, its performance is benchmarked against traditional approaches involving frequency band specific features and raw EEG signal features. This comparative analysis is intended to determine whether VMD offers a more robust representation of EEG activity.

Building upon these insights, this research proceeds to evaluate a range of regression models such as linear regression, elastic net, random forest, support vector regression (SVR), and extreme gradient boosting (XGBoost) to predict levels of consciousness based on resting-state EEG features. To address the inherent

limitations of individual models and to enhance predictive performance, this work proposes a novel ensemble of ensemble regressor model (EoERM) by integrating XGBoost, random forest, and SVR within a stacking-based ensemble learning framework. This research introduces a custom predictive equation for estimating the level of consciousness in individual subjects, which is normalized to align with the CRS-R score.

In parallel, this study employs advanced analytical approaches to enhance DOC classification based on functional connectivity (FC) measures namely phase lag index (PLI), weighted phase lag index (wPLI), debiased wPLI (dwPLI), amplitude envelope correlation (AEC), and coherence across standard EEG frequency bands. These connectivity-derived features serve as input for a series of deep learning models, including multilayer perceptrons (MLP), recurrent neural networks (RNN), long short-term memory (LSTM) networks, gated recurrent units (GRU), and convolutional neural network–LSTM (CNN-LSTM) hybrids. Furthermore, this study investigates the topological architecture of brain networks by constructing undirected weighted graphs based on dwPLI, coherence, and AEC. To analyze the clinical implications and explore the associations between graph-based features and the glasgow outcome scale–extended (GOSE), canonical correlation analysis (CCA) is employed.

## 1.2 Behavioural Assessment Methods

Behavioral assessment methods remain central in the evaluation of disorders of consciousness despite advancements in neuroimaging and computational analysis [12]. These assessments aim to gauge the awareness and responsiveness of DOC patients by observing and interpreting their behavioral responses to various stimuli. The most widely used tool for this purpose is the CRS-R score, which systematically examines auditory, visual, motor, oromotor, communication, and arousal responses [12]. CRS-R provides a structured approach to differentiate between DOC states

and is considered a gold standard due to its reliability and clinical relevance. By analyzing specific behaviors, such as purposeful eye movements, following commands, or vocalizations, CRS-R helps clinicians to identify signs of consciousness, even when minimal or inconsistent. While CRS-R has garnered widespread acceptance, other behavioral tools, like Glasgow coma scale (GCS) [18] and sensory modality assessment and rehabilitation technique (SMART) [19], also contribute to the evaluation of DOC patients. The GCS is often used in acute care settings as a quick method to assess consciousness by rating eye, verbal, and motor responses. However, its limitations in distinguishing nuanced DOC states, particularly MCS, have led to its supplementation with more detailed scales like CRS-R. SMART, on the other hand, is a more holistic assessment tool focusing on sensory responses and aims to identify subtle awareness markers over prolonged observation.

## 1.3 Resting State EEG Paradigms for DOC

Resting-state EEG has emerged as a crucial tool in the assessment and understanding of disorders of consciousness, offering a non-invasive and cost effective means of probing the brain's intrinsic neural activity [20]. Unlike task-based EEG paradigms, resting-state EEG does not require active patient participation, making it especially valuable for evaluating patients with limited or no behavioral responsiveness. This paradigm focuses on capturing spontaneous brain activity during a passive state, typically with the patient lying awake but not engaged in any specific task. The primary goal is to uncover patterns of neural activity and connectivity that reflect the level of consciousness and the integrity of underlying brain networks in DOC patients. Resting-state EEG provides insights into global and regional brain function, offering clues about the preservation or disruption of neural pathways critical to consciousness [11].

A key advantage of resting-state EEG is its ability to reveal the functional organization of the brain through measures such as spectral power, coherence, and

functional connectivity. Spectral power analysis examines the distribution of power across different frequency bands, including delta, theta, alpha, beta, and gamma waves, each associated with specific cognitive and physiological processes [21]. DOC patients often exhibit abnormal increase in delta power (indicative of a loss of higher cortical activity) and reduced alpha and beta activity, which are typically linked to wakefulness and cognition. These spectral abnormalities provide a window into the characteristic of altered brain states [14].

Functional connectivity analysis is another critical component of resting-state EEG paradigms, examining the synchronization of neural activity across different brain regions. Techniques such as phase locking value (PLV) [22], coherence, and Granger causality are used to quantify these connections [23]. Studies have demonstrated that DOC patients often show disrupted connectivity within and between major brain networks, particularly the default mode network (DMN), which plays a pivotal role in self referential thought and consciousness [24]. The extent of connectivity disruption offers valuable diagnostic and prognostic insight, as preserved or partially restored connectivity in the DMN, which is frequently associated with an increased likelihood of recovery from a DOC state.

Recent advancements in EEG analysis have introduced techniques like source localization and advanced machine learning models to enhance the diagnostic utility of resting-state EEG. Source localization methods, such as low-resolution electromagnetic tomography (LORETA) [25], can map EEG signals to their cortical origins, offering spatially resolved insights into brain activity. Machine learning algorithms further refine the analysis by identifying subtle patterns in resting-state EEG data that correlate with specific DOC states, enabling automated classification and prediction of recovery trajectories [26]. These tools help uncover covert consciousness, where traditional behavioral assessments may fail to detect signs of awareness.

## 1.4 Automated Classification of DOC Patients

Automated classification methods for DOC patients have transformed the landscape of clinical assessment, providing objective and precise tools to distinguish between different states of consciousness. Traditional classification of DOC relies on behavioral assessments, which are inherently subjective and prone to errors. Machine learning plays a central role in automated classification systems for DOC [20]. These methods involve training algorithms on labeled datasets of EEG signals to recognize patterns indicative of specific DOC states. Feature extraction is an important step in this process, where measurable attributes such as spectral power, functional connectivity, and entropy are derived from raw EEG data. Traditional machine learning models like support vector machines [26], random forests, and k-nearest neighbors (KNN) [27] have demonstrated success in classifying DOC patients by using such features. These models can also predict patient outcomes by analyzing longitudinal EEG data, offering valuable prognostic insights.

Deep learning, a subset of machine learning, has further enhanced automated classification methods by eliminating the need for manual feature engineering. Deep learning models, such as CNNs and RNNs, can learn hierarchical features directly from raw EEG signals, capturing spatial and temporal dependencies in the data [20]. CNNs are particularly effective at identifying spatial patterns across multichannel EEG data, while RNNs excel at modeling the temporal evolution of neural activity. These models have achieved remarkable accuracy in distinguishing between DOC states and detecting covert consciousness, where traditional behavioral assessments fail to identify signs of awareness [28]. Hybrid models combining CNNs and RNNs have been developed to integrate spatial and temporal features, further improving classification performance [29, 30].

## 1.5 Contribution of Thesis

This thesis represents a significant contribution to the field of disorders of consciousness study by addressing critical gaps in the diagnosis and classification of DOC patients through innovative methodologies and data-driven approaches. It also uncover the underlying neural characteristics associated with varying levels of consciousness. By utilizing advanced techniques, this thesis offers a deeper understanding of the subtle distinctions in brain activity across states, paving the way for more reliable diagnostic frameworks as an aid to the clinicians.

A major contribution is the generation of a robust and comprehensive resting-state EEG dataset specifically for DOC patients in collaboration with NIMHANS, Bangalore. This dataset, which includes high-quality recordings from diverse patient populations, serves as a valuable resource for future research. It not only addresses the pressing need for standardized, well labeled datasets but also provides the foundation for the development and validation of automated classification models.

Another significant aspect of the thesis is the application of variational mode decomposition as a preprocessing technique to enhance the classification ability of the classifiers implemented based on VMD-derived features. VMD, a powerful signal processing technique, decomposes EEG signals into intrinsic mode functions, enabling the identification of hidden neural patterns and frequency components associated with different consciousness levels. By leveraging VMD, this research enhances the accuracy of EEG signal analysis, providing a novel approach to classify DOC states based on their distinct neural signatures.

This thesis investigated the implementation and evaluation of various regression models for predicting consciousness levels in DOC patients using features derived from resting-state EEG. The models explored include linear regression, elastic net, random forest, SVR, and XGBoost. To overcome the limitations of individual models and to enhance predictive accuracy, a novel EoERM was introduced by integrating XGBoost, random forest, and SVR using a stacking-based ensemble

strategy. To enhance model transparency and facilitate interpretability, Shapley additive explanations (SHAP) were applied to each base regressor to identify the fifteen most important features contributing to the prediction. Furthermore, a custom consciousness scoring equation was developed by fitting the top 15 integrated features to a linear regression model, with the resulting coefficients serving as interpretable weights. This computed consciousness score were normalized to align with the CRS-R score, which ranges from 0 to 23, ensuring clinical compatibility and facilitating direct comparison with ground truth assessments.

Furthermore, the thesis explored the ability of functional connectivity measures in understanding the underlying neural correlation of different DOC cohorts. Firstly, functional connectivity were computed using various FC measures namely PLI, wPLI, dwPLI, AEC, and coherence across the five frequency bands. These features were used to train a series of neural network models, namely MLP, RNN, LSTM, GRU, and CNN-LSTM hybrid model. The comparative analysis demonstrated that AEC-based connectivity measure achieved better performance in classifying DOC subgroups. In parallel, the research also investigated the topological architecture of brain networks by constructing undirected weighted graphs from dwPLI, coherence, and AEC measures. Key graph theoretical metrics like clustering coefficient, path length, modularity, small-worldness, betweenness centrality, and node degree were extracted to assess network integration and segregation. Statistical analysis and CCA were employed to examine the relationship between graph-based features and the GOSE score.

## 1.6 Organization of Thesis

The main focus of this research is to classify disorders of consciousness patients using features extracted from resting-state EEG data, specifically recorded for this research. The research also aims to clinically validate the results, providing valuable support for clinicians in the diagnosis and classification of DOC subcategories.

Chapter 2 presents a comprehensive literature review focused on the use of EEG in the assessment and classification of disorders of consciousness. The chapter begins by detailing the advantages of resting-state EEG as a non-invasive and effective tool for studying spontaneous brain activity in non-responsive patients. It explores various EEG paradigms, including spectral analysis, non-linear dynamics, and functional connectivity, while highlighting their relevance in distinguishing between different states of consciousness. The chapter also examines the role of advanced analytical techniques, including machine learning and deep learning models, in enhancing diagnostic precision and prognosis prediction.

Chapter 3 details the construction of the resting state EEG dataset used in the study, in collaboration with NIMHANS, Bangalore for data acquisition. It outlines the inclusion criteria for DOC patients and healthy controls, and discusses the preprocessing pipeline adopted to prepare raw EEG signals for analysis. Techniques such as filtering, artifact removal, and normalization are described, along with the rationale for selecting specific preprocessing strategies.

Chapter 4 introduces variational mode decomposition as a preprocessing method in the EEG signal processing pipeline. This chapter elaborates on the use of VMD to generate features that encapsulate intrinsic signal dynamics, and compares these features to traditional approaches, such as frequency band-based and raw EEG-based metrics. Features including spectral entropy, sample entropy, kurtosis, and skewness are employed to capture a broad spectrum of neural activity. The chapter discusses the ability of VMD to distinguish DOC spectrum. It also compares the classification performance of VMD-derived features against traditional frequency band-based and raw EEG-based approaches using multiple machine learning classifiers, including an ensemble bagged tree model.

Chapter 5 presents the performance of various regressor models in predicting consciousness levels. Based on these findings, a novel ensemble of ensemble regressor model was proposed to predict the consciousness levels of individual subjects using

EEG-derived features. Further, a custom equation was introduced to predict consciousness levels utilizing the EoERM, both with and without feature selection through the SHAP method. The predicted consciousness score were normalized to the CRS-R score to validate the clinical implications of the proposed regressor model.

Chapter 6 focuses on the advanced analysis of resting-state EEG data to improve the classification of DOC patients by leveraging functional connectivity and graph theoretical frameworks. The functional connectivity measures applied include PLI, wPLI, dwPLI, AEC, and coherence across the five frequency bands. A series of neural network models were trained using the extracted features. Alongside, this chapter covers the construction of undirected weighted graphs from dwPLI, coherence, and AEC measures. Statistical analysis and CCA were employed to examine the relationship between graph-based features and the GOSE score.

Chapter 7 presents the conclusions and future directions of this research work.

# Chapter 2

# Literature Review

This chapter describes the literature relevant to the assessment of consciousness in patients with disorders of consciousness, focusing on behavioral methods, EEG analysis, and computational modeling techniques. The accurate assessment of consciousness in DOC patients remains a persistent clinical challenge, owing to the limitations of traditional behavioral evaluations and the complexity of neural data. While the CRS-R score is the widely adopted behavioral assessment method [12], they are inherently subjective and may fail to detect covert consciousness, leading to frequent misdiagnoses. In this context, resting-state EEG has emerged as a valuable, non-invasive tool for probing brain function, offering high temporal resolution and accessibility at the bedside [15]. However, the high dimensionality and variability of EEG signals compounded by inter-patient differences and diverse etiologies necessitate the application of advanced computational techniques for effective interpretation. In recent years, machine learning [31] and deep learning [20] approaches have shown promise in extracting meaningful patterns from EEG data, enabling the classification of consciousness states and supporting objective decision-making. This chapter reviews the literature across several domains: clinical and behavioral assessments of DOC, EEG-based biomarkers of consciousness, signal processing and feature extraction techniques, functional connectivity and graph theoretical analysis, and the application of explainable machine learning methods

in neurological research. Through this review, key gaps and limitations in current methodologies are identified, providing the foundation and motivation for the novel approaches proposed in this thesis.

## 2.1 Resting-state EEG Signal Paradigms for Diagnosis of DOC Patients

EEG has become a core neurophysiological modality in the study and clinical evaluation of DOC due to its non-invasive nature, affordability, bedside accessibility, and exceptional temporal resolution. Resting-state EEG has gained significant traction as a particularly advantageous paradigm for studying severely impaired or non-communicative patients [15]. It captures spontaneous brain activity without requiring any stimulus, thereby making it suitable for patients who cannot engage with external commands [20]. The different paradigms for analyzing resting-state EEG in disorders of consciousness are spectral and complexity measures, microstate analysis, functional connectivity, and graph theory [11]. The detailed description of the existing works utilizing these paradigms is given in the following section.

## 2.2 Spectral and Nonlinear Analysis of EEG

Spectral power analysis, a foundational aspect of frequency domain analysis, employs methods such as the Fourier transform to convert time domain EEG signals into their frequency domain representations [15]. This enables the examination of specific EEG frequency bands, including delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (>30 Hz), which have been shown to correlate with levels of consciousness. UWS patients typically exhibit elevated delta and theta activity, reflecting thalamo-cortical dysfunction, whereas MCS patients show relatively higher alpha power, which has been associated with improved clinical

prognosis [11, 32]. Several studies have employed non-linear dynamic measures to capture the inherent complexity of EEG signals [11]. Metrics such as Lempel-Ziv complexity, permutation entropy, and approximate entropy have been used to quantify the irregularity of brain activity [11]. Sitt et al. (2014) [33] combined multiple complexity and connectivity measures including weighted symbolic mutual information (wSMI), Kolmogorov-Chaitin complexity (KCC), permutation entropy (PE), and spectral entropy within an SVM-based framework, achieving an accuracy of 89% in differentiating MCS from UWS. However, this study did not include healthy controls or consider MCS subtypes, limiting its generalizability. Similarly, Engemann et al. (2015) [34] utilized PE and wSMI along with normalized alpha band power to discriminate between MCS and UWS patients using SVM classifiers, though specific performance metrics such as precision or recall were not reported. EEG microstate analysis, which segments EEG into quasi-stable topographical patterns, has also been employed to investigate consciousness. Reduced diversity in microstate types especially those related to alpha rhythms and a lower proportion of time spent in microstate D have been identified as markers distinguishing UWS from MCS [23, 35].

## 2.3 Functional Connectivity and Graph Theoretical Metrics

Functional connectivity analysis offers insight into the interaction between distributed brain regions by examining phase synchrony, coherence, and directed information flow. Measures such as coherence, imaginary coherence (IC) [23], PLI [33], and directed transfer function (DTF) [36] have been widely used to assess FC in DOC. Studies have shown that MCS patients exhibit stronger FC, particularly in the alpha and beta bands, compared to UWS patients [11]. Coherence-based metrics, while informative, often fall short in differentiating consciousness levels

due to volume conduction issues.

Alternative methods such as cross-approximate entropy and Granger causality have demonstrated more reliable discrimination by capturing the directional flow of information [23]. In particular, Granger causality has revealed asymmetries in incoming and outgoing connections in UWS patients [37]. Weighted symbolic mutual information has emerged as a robust, etiology independent FC metric. Dynamic functional connectivity (dFC) methods, which account for temporal variability in connectivity, have also shown promise and correlate well with behavioral scales like CRS-R.

Lehembre et al. (2012) [38] demonstrated the effectiveness of IC and PLI in distinguishing UWS from MCS, even using low-density EEG setups. Stefan et al. (2018) [23] further emphasized the diagnostic utility of alpha microstates and beta band clustering coefficients.

## 2.4 Graph Theory and Network Topology

Graph theoretical approaches provide a high-level framework for characterizing the topological structure of brain networks [39]. Key metrics such as global and local efficiency, modularity, and clustering coefficients have been used to quantify alterations in network architecture associated with DOC [17]. UWS patients typically exhibit decreased network efficiency and hub density compared to MCS patients, reflecting a breakdown in large-scale integration.

Naro et al. (2021) [32] applied multiplex and multilayer network analysis to resting-state EEG data from 17 UWS and 15 MCS patients. Their findings indicated degraded and heterogeneous functional networks, particularly within fronto-parietal circuits. However, these discriminative patterns were not evident when analyzing individual frequency bands in isolation.

Duclos et al. (2021) [40] used graph metrics and motif topology to distinguish levels of consciousness, noting greater global efficiency and participation coefficients

in MCS patients. Hao et al. (2024) [41] found that favorable outcomes were linked to lower clustering coefficients, higher participation, and greater network stability. Similar trends were observed by Sitt et al. (2014) [33] and Chennu et al. (2014) [14].

Rizkallah et al. (2019) [42] noted decreased participation coefficients and increased clustering in patients with lower consciousness levels, underscoring the relationship between network segregation and impaired awareness. Studies by Nadin et al. (2020) [43] and Kustermann et al. (2020) [44] further supported the prognostic value of graph theoretical features such as path length variance.

## 2.5 Advanced Signal Decomposition Techniques

Time-frequency analysis methods have become essential for interpreting non-stationary EEG signals. Wavelet transforms, empirical mode decomposition (EMD) [45], and VMD [46] enable signal decomposition into multiple components or modes. Wavelet-based techniques such as the adaptive flexible analytic wavelet transform, tunable quality factor wavelets [47], and rational dilation wavelets [48] have been applied across clinical and non-clinical domains, including Alzheimer's disease, ADHD, schizophrenia, emotion recognition, and motor imagery [49].

EMD and its extensions, including ensemble EMD (EEMD) [50], complementary EEMD [51], and CEEMDAN [52], address limitations such as mode mixing and sensitivity to noise. These methods yield robust intrinsic mode functions (IMFs), facilitating enhanced temporal and spectral resolution.

VMD, a more recent technique, formulates signal decomposition as a constrained variational problem. It introduces regularization to promote smoothness and reduce mode overlap, overcoming several drawbacks of EMD [53]. VMD has been successfully applied to EEG analysis in epilepsy, schizophrenia, Alzheimer's disease, ADHD, and emotion recognition [54–58]. It also finds utility in speech processing [59] and medical imaging.

## 2.6 Machine Learning and Deep Learning Applications

Machine learning (ML) and deep learning (DL) have been increasingly integrated into EEG analysis to support automated DOC diagnosis. Traditional ML models, including SVM, linear discriminant analysis (LDA), and ensemble classifiers, have shown promise using features such as spectral power, entropy, and FC metrics.

Di Gregorio et al. (2022) [31] used dominant frequency and mutual information features with LDA to classify DOC patients with traumatic and non-traumatic brain injuries, achieving 80–83.3% accuracy. Duclos et al. (2021) [40] evaluated wPLI and AEC in anesthetized subjects, reporting a maximum accuracy of 84.1% using SVM. Li et al. (2024) [24] achieved 96.2% accuracy using a CatBoost model with microstate and dynamic FC features. Despite high accuracy, concerns remain regarding the interpretability and standardization of such models.

The application of deep neural networks within the DOC domain remains relatively underexplored, likely due to the limited availability of patient data and EEG recordings, which constrains the feasibility of training deep learning models. A notable exception is the study by Craig et al. (2018) [60], which employed deep graph convolutional neural networks to model EEG data in this context.

Despite these limitations, emerging studies offer valuable insights into the potential directions for applying deep learning to EEG signal processing. Schirrmeister et al. (2017) [61] demonstrated that CNNs could achieve classification accuracies comparable to those obtained using filter bank common spatial pattern (FBCSP) methods when decoding task-related EEG signals. The authors developed visualization techniques to illustrate how CNNs leverage the amplitude of spectral band power features. A natural extension of this approach would involve visualizing the CNN's use of raw time domain EEG features, potentially offering insight into salient temporal patterns such as event related potentials (ERPs).

Zafar et al. (2017) [29] proposed a hybrid approach that combines CNNs with statistical methods. Specifically, they employed a t-test to select significant features prior to classification and utilized likelihood ratio-based score fusion to enhance predictive performance. Their algorithm, which processes multichannel EEG time series data, aligns with the principles of multivariate pattern analysis and highlights the efficacy of integrating deep learning with classical statistical techniques.

Similarly, Bashivan et al. (2015) [28] introduced a recurrent convolutional neural network architecture designed to preserve the spatial, spectral, and temporal properties of EEG signals. This framework aims to extract robust feature representations that are less sensitive to variability across sessions and subjects, as well as to the intrinsic noise commonly found in EEG recordings. The authors validated their method within the context of mental workload classification, demonstrating the advantages of their approach in capturing invariant patterns in multichannel EEG time series data.

This review highlights the diverse methodological landscape of EEG-based assessment in DOC, encompassing spectral, connectivity, non-linear, and machine learning approaches. While these techniques have advanced our understanding of the neural correlates of consciousness, significant challenges persist. These include model interpretability, dataset heterogeneity, and the need for standardized evaluation frameworks. Addressing these limitations is essential for translating research advances into clinically reliable diagnostic tools.

# Chapter 3

# EEG Signal Acquisition and Standardization

The resting-state EEG data of patients with disorders of consciousness were collected in accordance with the guidelines established by the American Academy of Neurology. The patient cohort included individuals with both traumatic and non-traumatic etiologies. These patients were admitted and evaluated as inpatients in the neurology wards and intensive care units of the neuroscience services at NIMHANS, Bangalore. The EEG signals of healthy control subjects were utilized as the reference standard for characterizing resting-state EEG features. Subjects were recruited based on the inclusion and exclusion criteria established by the clinician.

## 3.1 Experimental Setup

The study was conducted at the Departments of Neurology, Neuroanaesthesia, and Neurocritical Care at NIMHANS, Bangalore, India, employing a carefully designed experimental framework to ensure robust and reliable data collection. Figure 3.1 shows the EEG acquisition system used to record the resting state EEG of subjects in the wooden lab.



**Figure 3.1:** Resting-state EEG acquisition system

## 3.2 Inclusion/Exclusion Criteria

### Inclusion Criteria

- Adult patients (≥18 years) diagnosed with DOC (coma, vegetative state, or minimally conscious state) based on guidelines.
- Patients with DOC due to various etiologies (stroke, encephalitis, anoxia)
- Participants tested negative for COVID-19 (RT-PCR) and were admitted to hospital green zones (Both acute (≥7 days post-injury) and chronic cases were included).

### Exclusion Criteria

- Medically unstable patients where EEG was infeasible.
- Patients affected by progressive neurodegenerative disorders.
- Cases with acute reversible causes of DOC (≤7 days post-injury).
- Non-cooperative subjects or lack of informed consent.
- COVID-positive participants.

### Healthy Subjects

Healthy subjects were selected based on no history of brain injury, epilepsy, psychiatric, or neurodegenerative disorders. EEG recordings of healthy subjects from prior studies were utilized with appropriate permissions when necessary.

## 3.3 Details of the EEG System Used

EEG recordings were conducted using a standard 10-20 system, with a sampling rate of 512 Hz and open filter settings. Impedance levels were carefully maintained below 10 kΩ during the recording process to ensure data quality. Participants were in supine position with their eyes closed during the procedure. Reactivity was assessed through eye opening and closing responses for healthy subjects and responses to noxious stimuli and loud sounds for patients. The photic response was evaluated at frequencies ranging from 1 to 24 Hz with participants' eyes both open and closed. Efforts were made to obtain 30 minutes of artifact free EEG recordings whenever feasible. Figure 3.2 illustrates a subject in the supine position undergoing an EEG recording session.



**Figure 3.2:** Resting-state EEG recording setup

## 3.4 Demographics of the Subjects

Demographic data, including age, gender, presenting symptoms, and coma assessment scale scores, were systematically recorded. Comprehensive medical histories were documented, encompassing the site and extent of brain injury, imaging findings from computed tomography (CT) or magnetic resonance imaging (MRI), along with the details of prior treatments. EEG studies were performed on all consenting participants, with results analyzed and compared across the 3 groups. Recovery patterns were monitored through follow-up visits or telephonic interviews. While data collection was primarily conducted in a single session, additional sessions were arranged if the initial recordings were unusable. To ensure confidentiality, all data were anonymized, and only de-identified information was included in publications.

## 3.5 Quantitative and Qualitative Analysis of Data

Clinical evaluations of coma and behavioral states were conducted using standardized assessment scales, ensuring that evaluations were performed only when patients were neither under the influence of sedative medications nor in post-ictal states. Resting-state EEG data were systematically recorded and documented using a pre-designed proforma. Statistical analysis were carried out to compare findings between the patient group and healthy controls.

## 3.6 Ethical Clearance

The study received ethical approval from Institute Ethics Committee (BS&NS) at NIMHANS, Bangalore (No. NIMAHNS/$28^{th}$ IEC (BS & NS DIV.)/2020-21) and informed consent was obtained from all participants or their caregivers. No significant risks to participants were identified. EEG tests, conducted as part of the diagnostic protocols, were offered at no cost to patients categorized as below poverty line (BPL).

## 3.7 Sample Size Calculation

- Sample size of the dataset was calculated using MANOVA with G*Power software for 80% power and a small effect size.
- Data were collected from 20 subjects in each category of DOC (coma, UWS, and MCS) and from the healthy control group. Due to significant artifacts, four EEG recordings from the coma group were excluded. Tables 3.1, 3.2, and 3.3 show the detailed demographics of UWS, coma, and MCS groups, respectively.

**Table 3.1:** Demographics of the UWS group at baseline

| Subject ID | Age | Gender | Clinical Diagnosis | GCS | CRS-R |
|---|---|---|---|---|---|
| U1 | 54 | Female | Acute Disseminated Encephalomyelitis | 8 | 8 |
| U2 | 44 | Female | Encephalopathy | 11 | 10 |
| U3 | 44 | Male | Encephalopathy | 11 | 10 |
| U4 | 21 | Female | Encephalopathy | 6 | 5 |
| U5 | 36 | Female | Hypoxic Ischemic Encephalopathy | 6 | 7 |
| U6 | 26 | Female | Hypoxic Ischemic Encephalopathy | 9 | 9 |
| U7 | 42 | Male | Intracranial Haemorrhage | 11 | 8 |
| U8 | 34 | Female | Cerebral Venous Thrombosis | 7 | 4 |
| U9 | 36 | Male | Encephalopathy | 8 | 5 |
| U10 | 60 | Female | Encephalopathy | 8 | 5 |
| U11 | 67 | Female | Meningoencephalitis | 9 | 6 |
| U12 | 45 | Female | Meningoencephalitis | 8 | 7 |
| U13 | 63 | Male | Encephalopathy | 10 | 8 |
| U14 | 41 | Female | Metabolic /toxic encephalopathy | 10 | 10 |
| U15 | 33 | Female | Mass lesion | 7 | 5 |
| U16 | 19 | Male | Cerebrovascular | 8 | 6 |
| U17 | 26 | Male | Encephalopathy | 10 | 8 |
| U18 | 22 | Male | Auotoimmune- CNS Lupus | 13 | 18 |
| U19 | 39 | Male | Metabolic Encephalopathy | 10 | 9 |
| U20 | 32 | Male | Encephalopathy | 9 | 7 |

**Table 3.2:** Demographics of the coma group at baseline

| Subject ID | Age | Gender | Clinical Diagnosis | GCS | CRS-R |
|---|---|---|---|---|---|
| C1 | 91 | Male | Fronto parietal bleed | 3 | 0 |
| C2 | 26 | Female | TM with hydrocephalis | 3 | 1 |
| C3 | 75 | Male | Encephalopathy/posterior circulation seizure | 3 | 0 |
| C4 | 70 | Female | Neuroinfection | 4 | 1 |
| C5 | 40 | Male | Traumatic brain injury | 4 | 1 |
| C6 | 60 | Female | Stroke | 6 | 2 |
| C7 | 19 | Male | Bacterial meningitis | 3 | 1 |
| C8 | 23 | Male | Hypoxic encephalopathy | 3 | 0 |
| C9 | 25 | Male | Metabolic /toxic encephalopathy | 5 | 1 |
| C10 | 32 | Female | Post-operative/Hypoxic encephalopathy | 4 | 2 |
| C11 | 30 | Female | Meningoencephalitis | 6 | 2 |
| C12 | 28 | Female | Autoimmune-CNS lupus | 4 | 2 |
| C13 | 24 | Male | Meningioma | 4 | 1 |
| C14 | 66 | Male | Stroke | 3 | 1 |
| C15 | 57 | Male | Vestibular schwannoma | 3 | 0 |
| C16 | 26 | Female | TM with hydrocephalis | 3 | 0 |

**Table 3.3:** Demographics of the MCS group at baseline

| Subject ID | Age | Gender | Clinical Diagnosis | GCS | CRS-R |
|---|---|---|---|---|---|
| M1 | 36 | Male | Meningoencephalitis | 10 | 12 |
| M2 | 46 | Male | Tubercular Meninigitis | 11 | 14 |
| M3 | 40 | Female | TM with hydrocephalis | 11 | 15 |
| M4 | 75 | Male | Cerebral Venous Thrombosis | 7 | 10 |
| M5 | 31 | Male | Demyelination | 14 | 19 |
| M6 | 40 | Male | Intracranial Hemorrhage | 8 | 17 |
| M7 | 60 | Female | Intracranial Hemorrhage | 8 | 12 |
| M8 | 55 | Female | hypoxia | 11 | 12 |
| M9 | 45 | Female | Intracranial Hemorrhage | 12 | 16 |
| M10 | 40 | Female | Cerebral Venous Thrombosis | 8 | 16 |
| M11 | 30 | Female | Cerebral Venous Thrombosis | 9 | 15 |
| M12 | 55 | Male | Aneurysmal Intracranial Haemorrhage | 8 | 13 |
| M13 | 50 | Male | Meningoencephalitis | 10 | 11 |
| M14 | 75 | Male | Intracranial Hemorrhage | 5 | 11 |
| M15 | 79 | Female | Stroke | 10 | 17 |
| M16 | 64 | Male | Stroke | 14 | 18 |
| M17 | 43 | Male | Stroke | 7 | 12 |
| M18 | 70 | Male | Stroke | 14 | 18 |
| M19 | 25 | Male | Encephalopathy | 14 | 16 |
| M20 | 41 | Female | Encephalopathy | 12 | 13 |

This chapter detailed the procedures and protocols involved in the collection of EEG data for the study of disorders of consciousness. Key aspects included participant selection criteria, ethical considerations, recording settings, and preprocessing steps to ensure data quality and consistency. Special attention was given to minimizing artifacts and maintaining standardized conditions across sessions. The resulting dataset provides a robust foundation for subsequent analysis, enabling the application of various resting-state EEG paradigms to explore neural signatures of consciousness.

# Chapter 4

# Variational Mode Decomposition based EEG Signal Preprocessing and Classification by Machine Learning

The classification of disorders of consciousness presents significant challenges due to the complex, non-stationary nature of resting-state EEG signals. Traditional feature extraction methods often rely on predefined frequency bands, which may not fully capture the intricate temporal and spectral variations present in EEG data. Variational mode decomposition offers a robust alternative by decomposing signals into a set of intrinsic mode functions with improved mode separation and reduced mode mixing [46]. Unlike empirical mode decomposition [62], which suffers from issues such as sensitivity to noise and lack of backward error correction, VMD operates within a variational framework that enforces smoothness and adaptivity in signal decomposition [53]. This chapter explores the implementation of VMD as a preprocessing step for DOC classification, focusing on its potential to enhance feature extraction from EEG signals. By decomposing EEG data into well-defined modes, features such as sample entropy, spectral entropy, kurtosis, and skewness are

extracted and evaluated for their discriminative power. The statistical significance of these features is assessed using the Kruskal-Wallis test, ensuring that only the most relevant ones ($p < 0.05$) are utilized in subsequent classification experiments. The comparative analysis presented in this chapter examines the effectiveness of VMD-derived features against conventional frequency-based and raw EEG features, offering insights into its potential as a preprocessing method for improving DOC classification accuracy. The key contributions in this chapter are as follows:

- Proposes the use of VMD for EEG analysis in the classification of DOC patients, by integrating VMD with machine learning models to classify patients into coma, UWS, and MCS, thereby enhancing automated diagnostic accuracy.
- Employs the Kruskal-Wallis test to determine statistically significant features and refines classification models using only features with $p < 0.05$, leading to improved classification performance.
- Highlights the clinical implications of VMD-based feature extraction, emphasizing its potential to serve as a supportive supporting resource for behavioral assessment techniques and improve clinical decision-making in DOC diagnosis and treatment planning.

## 4.1 Feature Extraction

### 4.1.1 Spectral Power

EEG captures the brain's electrical activity through electrodes placed on the scalp, producing complex signals that reflect the underlying neural dynamics. One of the most widely used approaches to analyze these signals is spectral power analysis, which examines how the signal's energy is distributed across various frequency components over time [63]. This technique offers valuable insights into brain function by quantifying the strength of oscillatory activity within specific frequency

bands that are known to correspond to distinct physiological and cognitive states. EEG signals exhibit rhythmic patterns that can be categorized into five canonical frequency bands, each associated with different aspects of brain activity: delta ($\delta$), theta ($\theta$), alpha ($\alpha$), beta ($\beta$), gamma ($\gamma$) [64, 65].

To perform spectral analysis, the preprocessed EEG signal is first segmented and passed through bandpass filters tailored to each frequency band. The resulting frequency specific signals are then analyzed using power spectral density (PSD) estimation techniques to quantify the energy present within each band. In this study, PSD was computed using Welch's method, which involves segmenting the signal into overlapping windows of 4 second length and averaging the squared magnitude of the Fourier transform across these windows. This approach improves the reliability of the power estimation by reducing variance.

Spectral power can be expressed in two distinct ways, each offering unique interpretive value:

- Absolute Power (ab): Absolute power refers to the total amount of power contained within a specific frequency band, measured directly from the EEG signal without normalization. It represents the raw magnitude of oscillatory activity and is expressed in microvolts squared ($\mu V^2$) units. This measure is useful for assessing the intensity or overall strength of brain activity in a given frequency range, making it particularly valuable in clinical assessments where absolute changes in neural activity may be of interest.
- Relative Power (rp): Relative power, on the other hand, represents the proportion of power within a specific frequency band relative to the total power across all frequency bands. This normalized measure provides a more balanced view of how power is distributed throughout the spectrum, allowing for comparisons that are not influenced by individual differences in overall signal amplitude. It is calculated by dividing the PSD of the frequency band of interest by the total PSD summed across all bands and channels. This

proportional approach helps to highlight shifts in neural dynamics that may not be evident from absolute values alone, particularly in conditions involving altered states of consciousness or cognitive dysfunction.

By examining both absolute and relative spectral power, researchers can obtain a comprehensive picture of the brain's oscillatory behavior. This information can be used to investigate normal cognitive processes as well as pathological conditions, such as disorders of consciousness, providing critical data to support diagnosis, monitoring, and treatment planning.

### 4.1.2 Nonlinear dynamic features

Nonlinear dynamic features in EEG analysis pertain to the characterization of the brain's complex, nonlinear temporal behavior, which often cannot be adequately captured using conventional linear methods. Traditional linear analysis may overlook subtle and intricate patterns arising from the highly dynamic and interactive nature of neural activity. In contrast, nonlinear approaches, such as entropy-based measures, provide deeper insight into the underlying structure and irregularities of EEG signals, enabling the detection of variations that remain hidden in linear frameworks [11, 33].

Entropy, in this context, serves as a quantitative measure of signal's unpredictability or complexity. It reflects degree of randomness in a time series and is particularly useful for identifying changes in brain states or dysfunctions. In this study, two prominent entropy-based features are employed: permutation entropy and sample entropy [16]. These measures capture different aspects of signal complexity and are utilized to better understand the non-linear dynamics of EEG activity in relation to disorders of consciousness.

#### 4.1.2.1 Permutation Entropy

Permutation entropy is a widely used nonlinear dynamic feature employed to quantify the complexity of time series signals. Due to its computational efficiency and robustness, PE has been increasingly adopted for the analysis of biomedical signals, including EEG. The method involves transforming the original signal into a discrete symbol sequence, which captures the ordinal patterns within the signal. The complexity is then measured by estimating the probability distribution of these symbols in an m-dimensional embedding space [66].

Given an EEG signal $x(n)$ consisting of $N$ time samples, it can be embedded into a sequence of discrete symbols $s(t)$ based on two key parameters: the embedding dimension $D$, typically ranging from 3≤D≤7, and the embedding time delay $\tau$, which is set to 1 in this study [67]. From the embedded vectors, ordinal patterns are derived by ranking the amplitude values within each sub-sequence. These patterns are then mapped to $D!$ possible symbols, representing all permutations of the embedded sub-vectors [23, 67].

The resulting symbol sequence has a length of (N-(D-1)*$\tau$), and the probability of occurrence of each symbol $P_i$ is computed. Permutation entropy is then calculated using the Shannon entropy formula:

$$PE = -\sum_{i=1}^{m} P_i \log(P_i) \tag{4.1}$$

Where $P_i$ is the probability of occurrence of $i^{th}$ symbol.

Here, PE is computed for each channel across various frequency bands. The benefit of symbolizing the EEG signal lies in its ability to minimize sensitivity to noise while enhancing the efficiency of information quantification [67].

#### 4.1.2.2 Sample Entropy

Sample entropy (SampEn) is a widely used nonlinear metric designed to quantify the complexity and irregularity of time-series data such as EEG signals [68]. It serves as an improvement over approximate entropy by addressing several of its limitations. Specifically, sample entropy does not depend on the length of the time series, avoids self-matching, and offers a more straightforward and reliable implementation [69]. Due to these advantages, SampEn is considered a robust measure for analyzing complex physiological signals.

Mathematically, sample entropy is defined as the negative natural logarithm of the conditional probability that two sequences which are similar for $m$ consecutive data points remain similar at the next point $m+1$, within a given tolerance $r$. The statistic is denoted as $SampEn(m, r, N)$, where: $m$ is the embedding dimension or the length of sequences to be compared, $r$ is the tolerance window used to define similarity, $N$ is the length of the time series [68].

The computation of sample entropy involves determining the number of similar template matches of length $m$ and $m+1$, excluding self-matches. Let $x_m(i)$ be the embedded vector of length $m$ starting at point $i$. The number of vectors $x_m(j)$ within a distance $r$ from $x_m(i)$, excluding $i = j$, is denoted as $B_i$, and the corresponding average over all templates is given by:

$$B_i^m(r) = \frac{1}{(N-m-1)} B_i \tag{4.2}$$

where, $B_i$ is the number of vectors $x_m(j)$ at a distance of $x_m(i)$, without allowing self-counting, where j=1, N-m. Adding all the template vectors;

$$B^m(r) = \frac{1}{(N-m)} \sum_{i=1}^{N-m} B_i^m(r) \tag{4.3}$$

$A_i^m(r) = \frac{1}{(N-m-1)}$ [number of vectors $x_{m+1}^j$ at a distance of $x_(m+1)(i)$ ,without allowing self-counting, where [j=1, N-m].

Adding all the matches;

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r) \tag{4.4}$$

Where, $B^m(r)$ represents the probability that two sequences are similar for m points (possible), while $A^m(r)$ represents the probability that two sequences are similar for m+1 points (matches). From these equations, the conditional probability can be defined as; $\frac{A^m(r)}{B^m(r)}$ whose value is less than unity. Therefore the parameter sample entropy can be expressed as;

$$SampEn(m,r) = \lim_{N\to\infty} -\log \frac{A^m(r)}{B^m(r)} \tag{4.5}$$

which is derived from the static sample entropy,

$$SampEn(m,r,N) = -\log \frac{A^m(r)}{B^m(r)} \tag{4.6}$$

Features were extracted from all selected channels across the frequency bands delta, theta, alpha, and beta, as well as from the brain regions: frontal, central, occipital, parietal, and temporal, to create the feature vector. These feature vectors were then provided as input to various multi-class classifiers.

### 4.1.3 Spectral Entropy

Spectral entropy ($E_S$) is a frequency domain measure of signal complexity, derived as a normalized extension of Shannon entropy. It quantifies the irregularity or unpredictability of the power distribution across different frequency components within a given signal [70]. In the context of EEG analysis, spectral entropy assesses the degree of disorder in the frequency content of the signal, thereby serving as a valuable indicator of underlying neural dynamics [71].

To compute spectral entropy, the PSD of the EEG signal is first estimated using

methods such as Welch's technique. The power at each frequency is then normalized to form a probability distribution, $P(f)$, where $f$ denotes the frequency component. Shannon entropy formula is applied to this normalized power distribution to calculate spectral entropy as follows [71]:

$$E_S = -\sum(P(f) \cdot \log(P(f)) \tag{4.7}$$

Here, $P(f)$ represents the relative power at each frequency bin. The summation spans all frequency components within the specified band of interest.

Spectral entropy is sensitive to changes in the distribution of power across frequency bands, making it particularly suitable for applications involving altered states of consciousness. Higher values of $E_S$ correspond to more uniform or disordered power distributions, which are often associated with higher levels of awareness or consciousness. Lower spectral entropy values suggest more regular or concentrated frequency activity, commonly observed in states of reduced consciousness. Higher spectral entropy values in specific frequency bands signify a heightened state of consciousness, whereas lower values indicate a reduced level of consciousness.

### 4.1.4 Kurtosis

Kurtosis ($K$) is a statistical metric used to characterize the peakedness and tailedness of a probability distribution, offering valuable insights into the distributional shape of neural activity across various frequency components [72, 73]. In the context of EEG signal analysis, kurtosis serves as an indicator of how the amplitude distribution of a signal deviates from that of a Gaussian (normal) distribution [47].

Specifically, high kurtosis values reflect distributions that are sharply peaked with heavy tails, often indicating the presence of transient bursts or irregular fluctuations in neural activity. These characteristics are typically associated with reduced stability and increased signal variability. In contrast, low kurtosis values suggest a flatter distribution, indicative of more consistent and stable neural activity

over time [72].

The kurtosis of each IMF obtained via VMD is computed using the following statistical expression:

$$K = \frac{E[(x(n) - \mu)^4]}{[E[(x(n) - \mu)^2]]^2} \tag{4.8}$$

where $x(n)$ represents the signal, $\mu$ is the mean of the signal, and $E[\cdot]$ denotes the expected value.

Elevated kurtosis values within specific frequency bands or VMD modes may signal the presence of irregular or non-stationary brain dynamics, which have been linked to altered or diminished states of consciousness. Lower kurtosis values, as reflected by a more uniform signal distribution, are often associated with a more stable and conscious neural state [72].

### 4.1.5 Skewness

Skewness ($S_k$) is a higher-order statistical moment that quantifies the degree of asymmetry in a probability distribution. In the analysis of EEG signals, skewness provides insights into the imbalance or directional bias in neural activity across different frequency components [72, 73]. A positively skewed distribution implies that the data are concentrated toward the left of the mean, with a longer right tail, while a negatively skewed distribution suggests a concentration toward the right of the mean, with a longer left tail. Mathematically, skewness is defined as:

$$S_k = \frac{E[(x(n) - \mu)^3]}{\sigma^3} \tag{4.9}$$

In the context of consciousness assessment, elevated skewness values within specific frequency bands or decomposed modes may indicate an asymmetric distribution of neural activity, often associated with abnormal or disrupted brain function, potentially reflecting a diminished level of consciousness. In contrast, lower skewness values, representing a more symmetric distribution of signal amplitudes, are

generally indicative of stable and balanced neural activity, characteristic of more conscious brain states [73].

## 4.2 Variational Mode Decomposition

Variational mode decomposition is an adaptive, non-recursive signal decomposition technique that decomposes a complex signal into a set of band limited IMFs, each with a specific center frequency and narrow bandwidth [46]. Introduced by Dragomiretskiy and Zosso in 2014 [46], VMD is formulated within a variational framework, setting it apart from classical techniques such as EMD and CEEMD. While EMD relies on iterative sifting and is prone to mode mixing and sensitivity to noise, VMD overcomes these challenges by adopting a more rigorous mathematical foundation and by offering enhanced noise robustness, convergence guarantees, and fixed decomposition.

VMD has become a powerful tool for analyzing non-stationary and nonlinear signals, such as EEG, where conventional fourier-based methods fail to capture localized and transient oscillatory components effectively. EEG signals, particularly during resting states, exhibit complex oscillatory patterns that vary across time and frequency domains. The decomposition of these signals into informative modes facilitates more accurate feature extraction and subsequent classification, especially in disorders of consciousness and other clinical applications.

VMD seeks to decompose a signal $f(t)$ into $k$ intrinsic mode functions $\{u_k(t)\}$ each with its corresponding center frequency $\omega_k$. The decomposition problem is formulated as a constrained variational optimization problem. The decomposition of the input signal is devised such that the sum of these modes reconstructs the original signal as accurately as possible. The core assumptions include each mode is amplitude and frequency modulated, each mode is mostly compact in the frequency domain (i.e., has limited bandwidth), and each mode oscillates around its own center frequency. The decomposition problem is cast as a constrained optimization

task:

$$J = \sum_{k=1}^{K} \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] \; e^{-j\omega_k t} \right\|_2^2$$

The minimization problem $\min(u_k,\, \omega_k)\{J\}$ is solved by locating the saddle point of augmented Lagrangian $\mathcal{L}$ using alternate direction method of multipliers. This is defined as,

$$\mathcal{L}(u_k, \omega_k, \lambda) = \alpha \sum_{k=1}^{K} \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] \; e^{-j\omega_k t} \right\|_2^2 + \left\| x(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2$$

$$+ \left\langle \lambda(t), x(t) - \sum_{k=1}^{K} u_k(t) \right\rangle \tag{4.10}$$

$$\sum_{k=1}^{K} u_k(t) = f(t) \tag{4.11}$$

Here:

- $*$ denotes convolution.
- $\partial_t$ is the partial derivative with respect to time.
- $\delta(t) + \frac{j}{\pi t}$ is the Hilbert transform kernel.

This formulation minimizes the sum of bandwidths of the modes in the frequency domain, where:

The analytic signal is obtained using the Hilbert transform,

Frequency shifting to baseband is done via complex exponential The bandwidth is quantified by the squared norm of the gradient of the demodulated analytic signal. To solve the constrained optimization, the method introduces an augmented Lagrangian:

Where:

- $\lambda(t)$ is the Lagrange multiplier,
- $\alpha$ is the penalty factor for the constraint,
- $\langle \cdot, \cdot \rangle$ denotes the inner product.

### VMD Iterative Solution (in Frequency Domain)

The optimization is carried out using the Alternating Direction Method of Multipliers (ADMM), which updates each mode $\hat{u}_k(\omega)$ and its center frequency $\omega_k$ in the Fourier domain iteratively.

**(a) Update of $\hat{u}_k(\omega)$:**

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \tag{4.12}$$

**(b) Update of center frequency $\omega_k$:**

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \tag{4.13}$$

**(c) Update of Lagrange multiplier:**

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau \left( \hat{f}(\omega) - \sum_{k=1}^{K} \hat{u}_k(\omega) \right) \tag{4.14}$$

Here:

- $\hat{u}_k(\omega)$ and $\hat{f}(\omega)$ are the Fourier transforms of $u_k(t)$ and $f(t)$,
- $\tau$ is the update rate for the Lagrange multiplier.

## 4.3 Methodology

This work tackles the challenges of classifying patients with disorders of consciousness using resting-state EEG signals, which are complex and non-stationary. To address these difficulties, the study proposes applying VMD to preprocessed EEG signals before feature extraction and machine learning classification. VMD is used to decompose the raw EEG signals into distinct intrinsic modes, allowing the separation of different frequency components. This decomposition serves as a preprocessing step to enhance the subsequent feature extraction process. After

decomposition, features such as sample entropy, spectral entropy, kurtosis, and skewness are calculated for each of the modes, capturing a wide range of signal characteristics that can aid in classification. These features are then compared with those extracted using traditional frequency band-based approaches and directly from raw EEG data, providing a comprehensive evaluation of the effectiveness of VMD in improving classification performance. The classification task involves two key processes: binary classification, distinguishing between UWS and MCS, and multi-class classification, which categorizes patients into three states: coma, UWS, and MCS. The classification performance is evaluated by applying machine learning algorithms to the features extracted from the different preprocessing approaches. Statistical analysis is performed using the Kruskal-Wallis test to assess the significance of the extracted features, with those showing a p-value of less than 0.05 being selected for further analysis and classification. This ensures that only the most relevant and statistically significant features are used in the final classification models. The proposed methodology is illustrated in Figure 4.1, which outlines the sequence of steps involved in the approach and the experiments conducted. The results underscore the significance of using VMD for signal decomposition as a preprocessing step in improving the accuracy of machine learning-based classification, with implications for more reliable diagnosis and informed decision-making in the treatment of patients with disorders of consciousness.

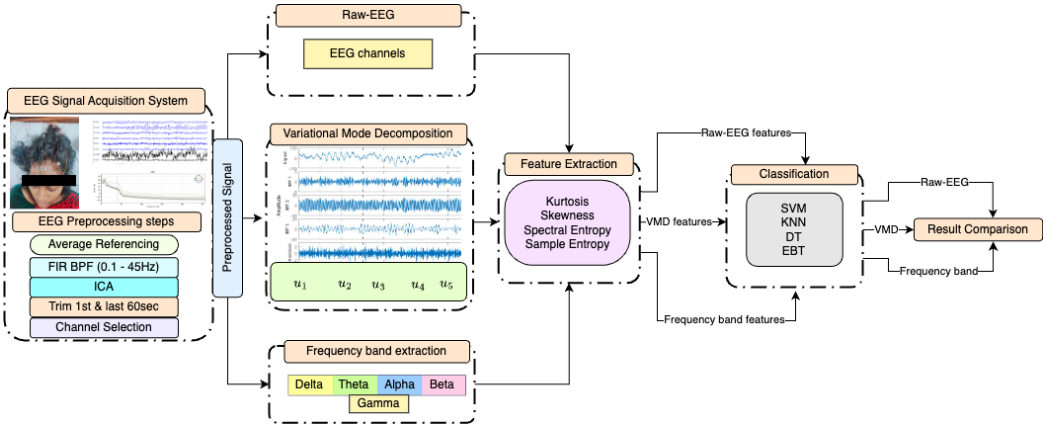



**Figure 4.1:** Block diagram of proposed approach.

## 4.4 Classifier Models

For classifying the DOC cohort based on VMD mode-based, frequency band-based, and raw EEG-based features, two types of classifications are carried out. In one, binary classification is performed between the UWS and MCS categories. In the other, multi-class classification is carried out among the coma, UWS, and MCS categories. For binary classification, the dataset consisted of 30 patients, whereas for multiclass classification, it included 45 patients. The selection of appropriate models for a machine learning task must be guided by both the characteristics of the data and the specific objectives of the study. Given the challenges associated with acquiring EEG data from patients in a state of disorders of consciousness, publicly available databases with a sufficiently large patient sample are lacking, which limits the use of advanced classifiers such as end-to-end deep learning models. Conventional machine learning models are employed to address this task. The desired classifier model must be capable of handling only two or three classes.

Multiple classifiers [74] were employed to assess their performances across different approaches and identify the most effective one for the task at hand. Decision trees (DT) are appropriate for classification problems where the training data may contain errors or noise, and thus chose them as one possible classifier. The KNN algorithm performs effectively for both binary and multi-class classification tasks, making it another suitable choice for this research. Support vector machine (SVM) is a powerful machine learning algorithm that has been used for a variety of classification tasks. SVMs are efficient and adaptable since they can handle high-dimensional data and nonlinear relationships. SVM is very effective since it finds the maximum separating hyperplane between the different classes in the feature space. So, SVM is selected as the third classifier. An ensemble bagged tree (EBT) classifier aggregates the predictions from multiple decision trees, each trained on random subsets of the data and features. This approach mitigates the risk of overfitting and enhances the model's generalization capabilities. Therefore, EBT

was also included as a classifier. In total, four distinct classifiers were employed, and their performance was evaluated using several metrics, including accuracy, precision, recall, and F1-score.

## 4.5 Results

This research investigated and compared the performance of features derived from five VMD modes, EEG frequency bands (delta, theta, alpha, beta, and gamma), and the 17 selected channels for analysis. The extracted features include spectral entropy, sample entropy, kurtosis, and skewness. Data analysis and classification were performed using Python. Prior to extracting VMD modes and frequency bands, the feature variations between groups were examined for each channel. The classification of disorders of consciousness groups was conducted, encompassing both binary and multi-class scenarios. To mitigate the risk of overfitting, a 10-fold cross-validation procedure was implemented.

### 4.5.1 Approach 1: Features Derived from the Raw EEG

Features were extracted from the preprocessed EEG signals recorded across 17 channels, yielding a total of 68 features (17 channels $\times$ 4 features per channel). These features were analyzed to investigate their variations across the three patient groups: coma, UWS, and MCS. The analysis revealed that patients in the MCS group exhibited EEG signals with higher complexity, suggesting comparatively greater levels of awareness and wakefulness than those observed in the coma and UWS groups.

The inter-group variations in the mean values of sample entropy and spectral entropy across all channels are illustrated in Figure 4.2 and Figure 4.3, respectively. In a similar manner, Figure 4.4 and Figure 4.5 depict the differences in mean kurtosis and skewness values for each channel among the three groups.

These extracted features from each EEG channel were utilized as input for classification experiments, encompassing both binary and multi-class classification tasks. The results of the binary classification between UWS and MCS patients are presented in Table 4.1. Among the classifiers evaluated, the linear support vector machine (Li.SVM) demonstrated superior performance, achieving the highest classification accuracy of 73.3%. Li.SVM yielded higher values for precision, recall, and F1-score, underscoring its effectiveness in discriminating between UWS and MCS states.

Table 4.2 summarizes the performance of various classifiers in the multi-class classification task involving coma, UWS, and MCS patients. Among the models assessed, the ensemble of bagged decision trees and the KNN classifier outperformed the other algorithms, attaining classification accuracies of 45.5% and 46%, respectively. However, these results suggest that distinguishing among the three consciousness states remains a challenging task, particularly in the presence of overlapping EEG features.

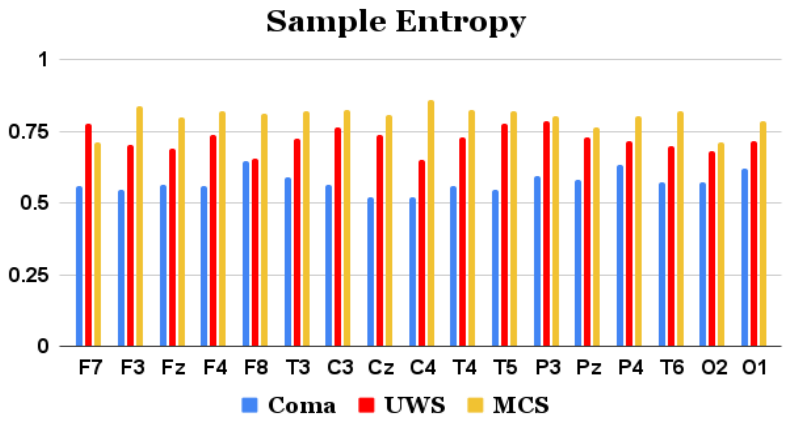



**Figure 4.2:** Mean sample entropy values

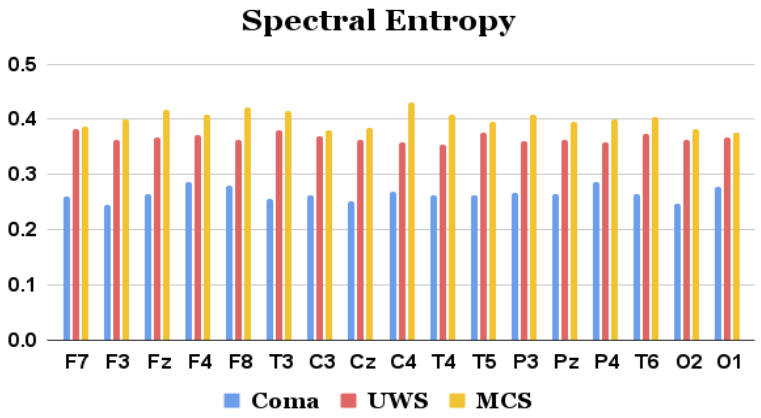



**Figure 4.3:** Mean spectral entropy values

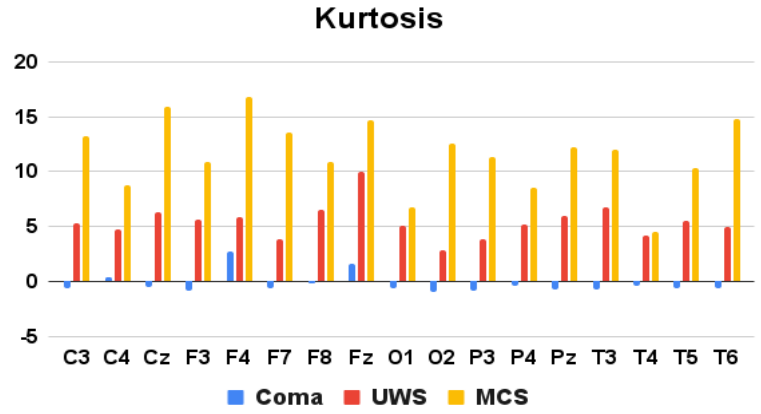



**Figure 4.4:** Mean kurtosis values

**Table 4.1:** Binary (UWS vs MCS) classification results (in %) using 10-fold cross-validation with all four features directly extracted from the 17 channels of the raw EEG. Number of patients: 30

| Classifier | Accuracy | Precision | Recall | F1 score |
|---|---|---|---|---|
| KNN | 56.7 | 45 | 55 | 49.5 |
| Li.SVM | 73.3 | 65.8 | 72.5 | 69 |
| DT | 53.3 | 47.5 | 52.5 | 49.9 |
| EBT | 60 | 55.8 | 62.5 | 59 |

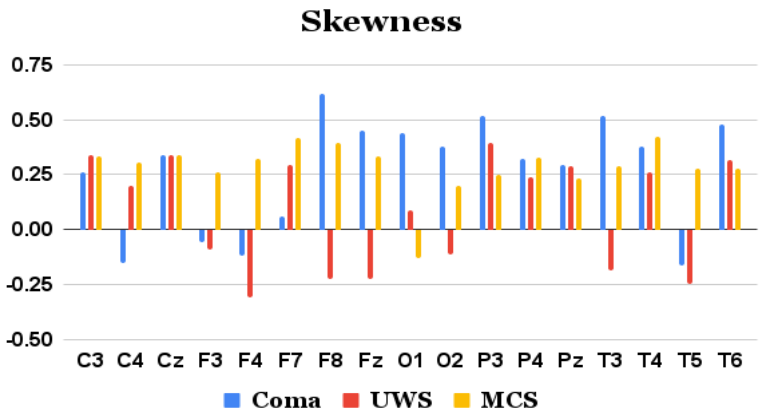



**Figure 4.5:** Mean skewness values

**Table 4.2:** Multi-class (Coma vs UWS vs MCS) classification results (in %) using 10-fold cross-validation with all the four features extracted from the 17 channels of the raw EEG. Number of patients: 45

| Classifier | Accuracy | Precision | Recall | F1 score |
|---|---|---|---|---|
| KNN | 46 | 37.2 | 48.3 | 42 |
| Li.SVM | 35.5 | 28.3 | 40 | 33.1 |
| DT | 40 | 31.8 | 41.7 | 36.1 |
| EBT | 45.5 | 40.7 | 48.3 | 44.2 |

### 4.5.2 Approach 2: Features Derived from the Frequency Bands

The five canonical frequency bands of the EEG signal: delta, theta, alpha, beta, and gamma were extracted from the preprocessed signals using a zero-phase, non-causal finite impulse response (FIR) bandpass filter. The filter design was based on a windowed time domain approach (implemented using the firwin function), with band specific lower and upper cut-off frequencies. A Hamming window was employed, offering a passband ripple of 0.0194 and a stopband attenuation of 53 dB, to ensure minimal signal distortion and adequate suppression of out-of-band frequencies.

Following bandpass filtering, four statistical and entropy-based features were extracted from each of the five frequency bands across all 17 EEG channels. This

resulted in a total of 340 features (17 channels $\times$ 5 frequency bands $\times$ 4 features), which were used as inputs to multiple machine learning classifiers.

The performance of these frequency band-derived features in the binary classification of UWS versus MCS is presented in Table 4.3. Among the classifiers evaluated, the KNN algorithm demonstrated relatively poor performance, achieving a classification accuracy of only 36.7%. In contrast, the Li.SVM yielded a significantly higher accuracy of 66.7%, indicating superior discriminative ability. Decision tree and ensemble of bagged trees classifiers performed less effectively, with accuracies of 30% and 36.7%, respectively.

Detailed performance metrics further highlight these trends. KNN achieved a precision of 35.8%, recall of 42.5%, and F1-score of 38.5%, indicating limited capability in accurately identifying and classifying positive instances. Li.SVM, on the other hand, demonstrated more robust results with a precision of 58.3%, recall of 75%, and an F1-score of 65.6%. The DT classifier exhibited particularly low performance, with precision, recall, and F1-score values of 26.7%, 30%, and 28.3%, respectively. Similarly, EBT attained a precision of 25.8%, recall of 37.5%, and an F1-score of 30.6%, reflecting its limited effectiveness in this binary classification task.

The multi-class classification performance of the same classifiers in distinguishing among coma, UWS, and MCS patient groups is summarized in Table 4.4. Among the models assessed, the linear support vector machine again demonstrated relatively better performance compared to the other classifiers. Nevertheless, the overall accuracy remained modest, with a maximum of only 45%, underscoring the inherent difficulty of multi-class classification in the presence of overlapping EEG features and subtle distinctions in consciousness states.

**Table 4.3:** Binary (UWS vs MCS) classification results (in %) using 10-fold cross-validation with the entire set of frequency band-based features. Number of patients: 30. The performance is inadequate with all the classifiers

| Classifier | Accuracy | Precision | Recall | F1 score |
|---|---|---|---|---|
| KNN | 36.7 | 35.8 | 42.5 | 38.5 |
| Li.SVM | 66.7 | 58.3 | 75 | 65.6 |
| DT | 30 | 26.7 | 30 | 28.3 |
| EBT | 36.7 | 25.8 | 37.5 | 30.6 |

**Table 4.4:** Multi-class (Coma vs UWS vs MCS) classification results (in %) using 10-fold cross-validation for the entire set of frequency band-derived features. Number of patients: 45. The performance is unsatisfactory with all the classifiers

| Classifier | Accuracy | Precision | Recall | F1 score |
|---|---|---|---|---|
| KNN | 34.5 | 32.5 | 35 | 33.7 |
| Li.SVM | 45 | 36.1 | 41.7 | 38.7 |
| DT | 43.5 | 36.1 | 43.3 | 39.4 |
| EBT | 42.5 | 34.2 | 45 | 38.9 |

### 4.5.3 Approach 3: Features Derived from the VMD Modes

In this approach, the same four features namely, sample entropy, spectral entropy, kurtosis, and skewness were extracted from each of the five IMFs obtained through VMD, across all 17 EEG channels. This resulted in a total of 340 features (17 channels $\times$ 5 modes $\times$ 4 features). The VMD parameters were configured as follows: the penalty factor $\alpha$ was set to 2000, the noise tolerance parameter $\tau$ was assigned a value of 0, and the center frequencies (omegas) were initialized uniformly with a value of 1. The extracted VMD-based features were used as input to various classification algorithms to evaluate their discriminative performance.

Table 4.5 presents the binary classification results distinguishing between UWS and MCS patients. The KNN classifier correctly identified 60% of the instances. However, its precision and recall were relatively low, at 44.2% and 57.5%, respectively, resulting in an F1-score of 50%, which reflects limited ability to reliably detect both positive and negative classes.

The SVM classifier with a linear kernel demonstrated markedly superior performance, achieving an accuracy of 80%. The DT classifier yielded a moderate accuracy of 63.3%. Among the classifiers evaluated, the EBT demonstrated the highest classification performance, with an accuracy of 83.3%. EBT classifier achieved a precision of 84.2%, a recall of 87.5%, and an F1-score of 85.8%, indicating its robust capability in accurately distinguishing between the UWS and MCS states.

**Table 4.5:** Binary (UWS vs MCS) classification results (in %) using 10-fold cross-validation employing all the VMD-based features. Number of patients: 30. Ensemble bagged tree performs the best with VMD-based features

| **Classifier** | **Accuracy** | **Precision** | **Recall** | **F1 score** |
|---|---|---|---|---|
| KNN | 60 | 44.2 | 57.5 | 50 |
| Li.SVM | 80 | 77.5 | 82.5 | 79.9 |
| DT | 63.3 | 62.5 | 67.5 | 64.9 |
| **EBT** | **83.3** | **84.2** | **87.5** | **85.8** |

Table 4.6 presents the results of the multi-class classification task involving coma, UWS, and MCS patients, using features derived from VMD. The performance of several classifiers was evaluated based on accuracy, precision, recall, and F1 score.

The k-nearest neighbor classifier demonstrated limited effectiveness, achieving an accuracy of only 46%. It recorded a precision of 32.3% and a recall of 45%, indicating a relatively poor capacity to correctly classify and identify true instances across the three patient groups.

The support vector machine classifier with a linear kernel showed moderate improvement, reaching an accuracy of 57%. However, it was outperformed by the DT classifier, which attained an accuracy of 73.5%. The DT also achieved a precision of 67.2% and a recall of 73.3%, reflecting its reasonable ability to identify true positive cases across the classes. The resulting F1 score of 70.1% indicates a balanced performance between precision and recall.

Among all classifiers evaluated, the EBT demonstrated the best performance, achieving the highest classification accuracy of 76%. It also recorded a precision of 78.9%, suggesting a low rate of false positives, and a recall of 78.3%, indicating strong capability in identifying true instances across the three diagnostic categories. The F1 score of 78.6% further supports the classifier's overall balanced and superior classification performance in the multi-class setting.

**Table 4.6:** Multi-class (Coma vs UWS vs MCS) classification results (in %) using 10-fold cross-validation employing all the VMD-based features. Number of patients: 45. Ensemble bagged tree performs the best with VMD-based features

| Classifier | Accuracy | Precision | Recall | F1 score |
|---|---|---|---|---|
| KNN | 46 | 32.3 | 45 | 37.6 |
| Li.SVM | 57 | 46.9 | 56.7 | 51.3 |
| DT | 73.5 | 67.2 | 73.3 | 70.1 |
| **EBT** | **76** | **78.9** | **78.3** | **78.6** |

### 4.5.4 Performance Comparison Between the Three Approaches

Table 4.7 summarizes the comparative improvement in classification accuracy achieved through the use of features derived from VMD, as opposed to those extracted from conventional frequency band decomposition and raw EEG signals. Across all classifiers excluding the KNN classifier in the multi-class setting the VMD-based feature extraction method consistently outperformed the other two approaches.

In particular, the EBT classifier demonstrated a notable enhancement in classification performance. For multi-class classification, the accuracy increased from 42.5% (achieved using frequency band-derived features) to 76% with VMD-derived features, representing a relative improvement of approximately 78.8%. Likewise, for binary classification between UWS and MCS states, the accuracy improved from 36.7% to 83.3%, yielding a substantial relative improvement of 127%. These

results underscore the efficacy of VMD as a robust feature extraction method, contributing to significantly enhanced discriminative power in both binary and multi-class classification tasks related to the assessment of consciousness levels.

**Table 4.7:** Comparison of percentage increase in accuracy of VMD-derived features over frequency band-derived (FB) and raw EEG-derived (EEG) features, when all the features are employed in each of the approaches

| | Multiclass classification | | Binary classification | |
|---|---|---|---|---|
| Classifier | FB-features | Raw EEG-features | FB-features | Raw EEG-features |
| KNN | 33.3 | 0 | 63.5 | 5.8 |
| Li.SVM | 26.7 | 60.6 | 19.9 | 9.1 |
| DT | 69 | 83.8 | 111 | 18.8 |
| EBT | 78.8 | 67.0 | 127 | 38.8 |

## 4.6 Statistical Analysis

Statistical analysis of the extracted EEG features was performed using the Kruskal–Wallis test, selected due to its suitability for the nature of the data and the study design. The Kruskal–Wallis test is a non-parametric, rank-based method used to compare the medians of two or more independent groups [75]. Unlike one-way ANOVA, it does not assume normally distributed data, making it particularly appropriate for datasets such as EEG signals, which are often characterized by non-normal distributions, high variability, and the presence of outliers.

The choice of the Kruskal–Wallis test was informed by several advantageous properties:

- **Non-parametric**: The method does not require the data to follow a Gaussian distribution, making it well-suited for skewed or ordinal data.
- **Robustness to Outliers**: Because the test is based on ranks rather than raw data values, it is less sensitive to extreme values, providing more reliable results in the presence of outliers.

- **Applicability to Multiple Groups**: Unlike the Mann–Whitney U test, which is restricted to two groups, the Kruskal–Wallis test can be applied to more than two independent groups, offering greater flexibility in multi-group comparisons.

These characteristics render the Kruskal–Wallis test particularly appropriate for EEG signal analysis. However, while the test can determine whether statistically significant differences exist among groups, it does not indicate which specific groups differ. Therefore, post-hoc analysis are required for detailed group-wise comparisons. The test assumes that the distributions of the groups being compared have a similar shape; violations of this assumption may affect the validity of the results.

Here, the Kruskal–Wallis test was used to compare three patient groups: coma, UWS), and MCS. The test was applied independently to each feature derived from three different preprocessing approaches: raw EEG signals, frequency band decomposition, and VMD. A significance threshold of $p < 0.05$ was used to determine statistical significance.

Following the Kruskal–Wallis analysis, Dunn's post-hoc test was employed to identify the specific pairs of groups that exhibited statistically significant differences. Features that achieved significance ($p < 0.05$) were selected from each of the three approaches to evaluate their potential in enhancing classification performance. Both binary and multi-class classification tasks were repeated using only the statistically significant features. These classification experiments were conducted using 10-fold cross-validation to ensure robust and reliable performance evaluation.

### 4.6.1 Clinical Implications of Correlations

The Spearman rank correlation method was used to estimate the correlation coefficients between the significant feature-electrode combinations extracted from the raw EEG approach, the frequency band-based approach, and the VMD approach with the CRS-R and GCS scores. The correlation coefficients of each of the chosen

electrode feature combinations for the different approaches are shown in Figures 4.6 to 4.11. It can be seen that feature selection is effective in that most of the chosen features have correlation coefficients with absolute values varying between 0.2 and 0.5.

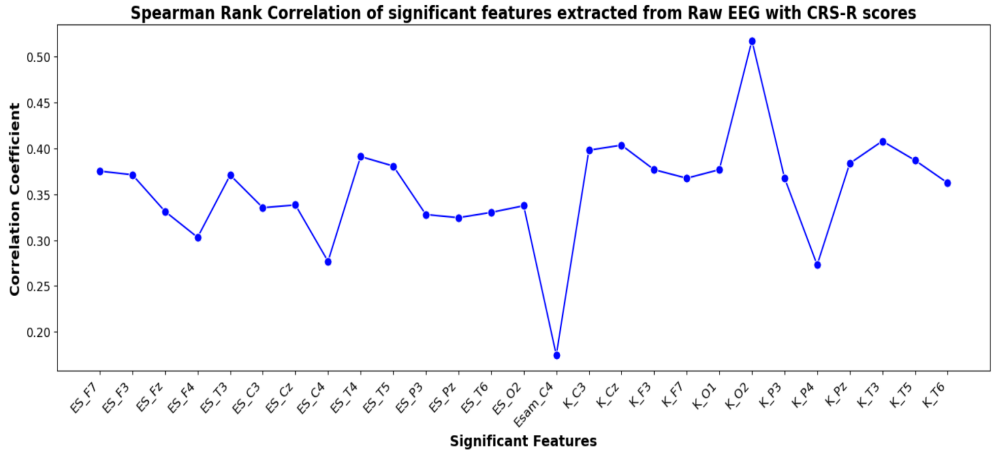



**Figure 4.6:** Spearman rank correlation coefficients between the CRS-R scores and selected features extracted from each of the 17 EEG channels

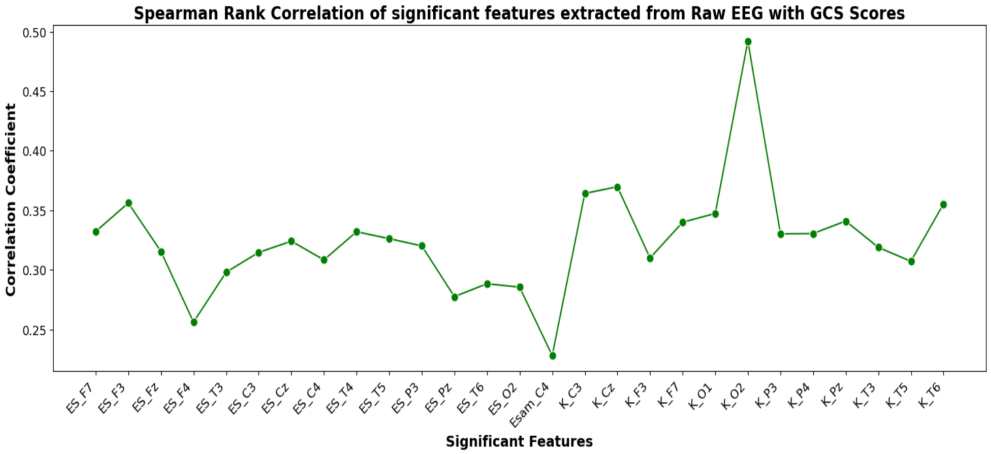



**Figure 4.7:** Spearman rank correlation coefficients between the GCS scores and selected features extracted from each of the 17 EEG channels

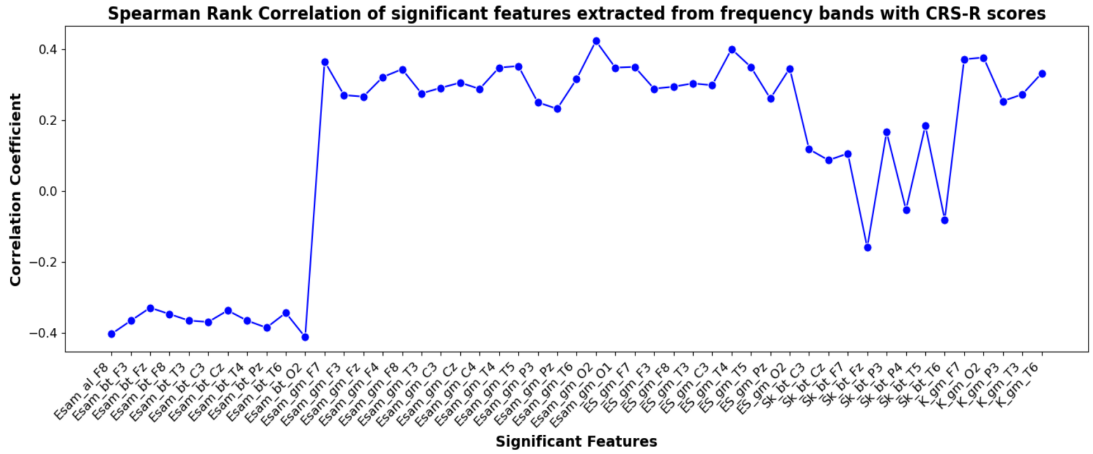



**Figure 4.8:** Spearman rank correlation coefficients between the CRS-R scores and selected features extracted from each frequency bands

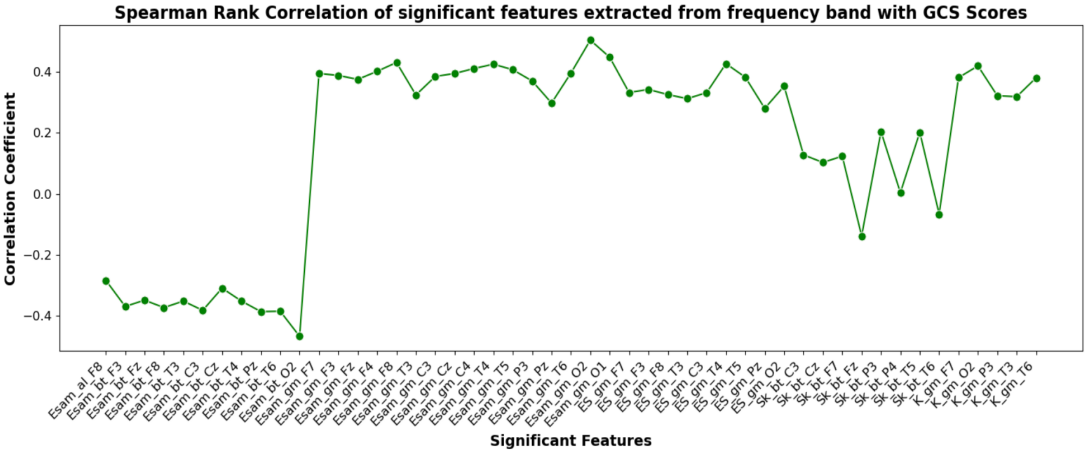



**Figure 4.9:** Spearman rank correlation coefficients between the GCS scores and selected features extracted from each frequency bands

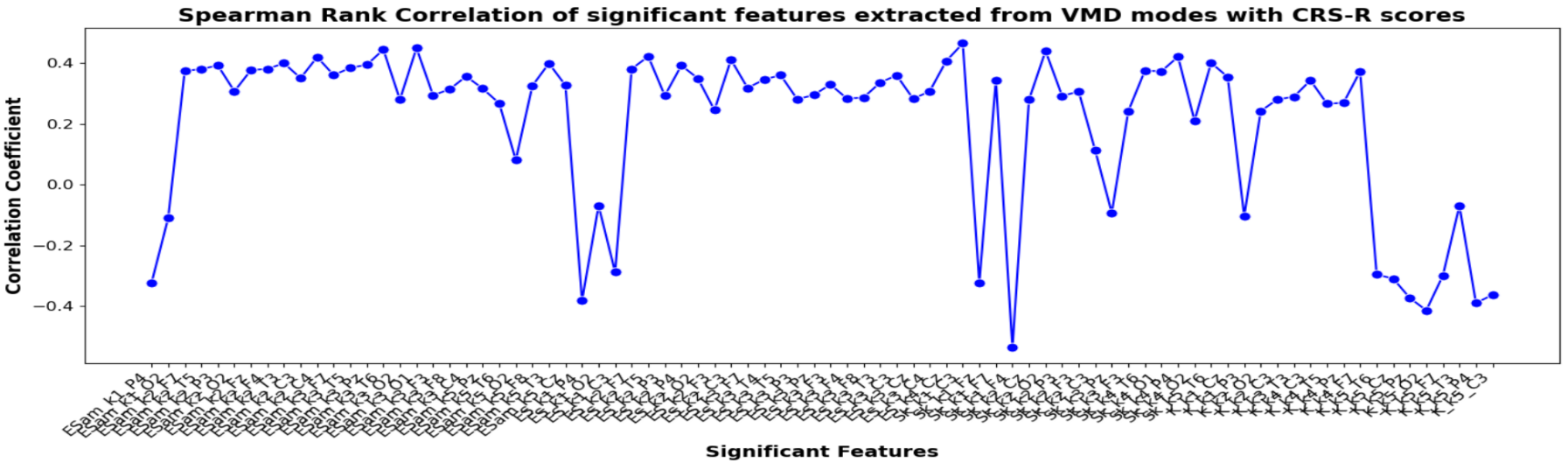



**Figure 4.10:** Spearman rank correlation coefficients between the CRS-R scores and selected features extracted from each of the five VMD modes

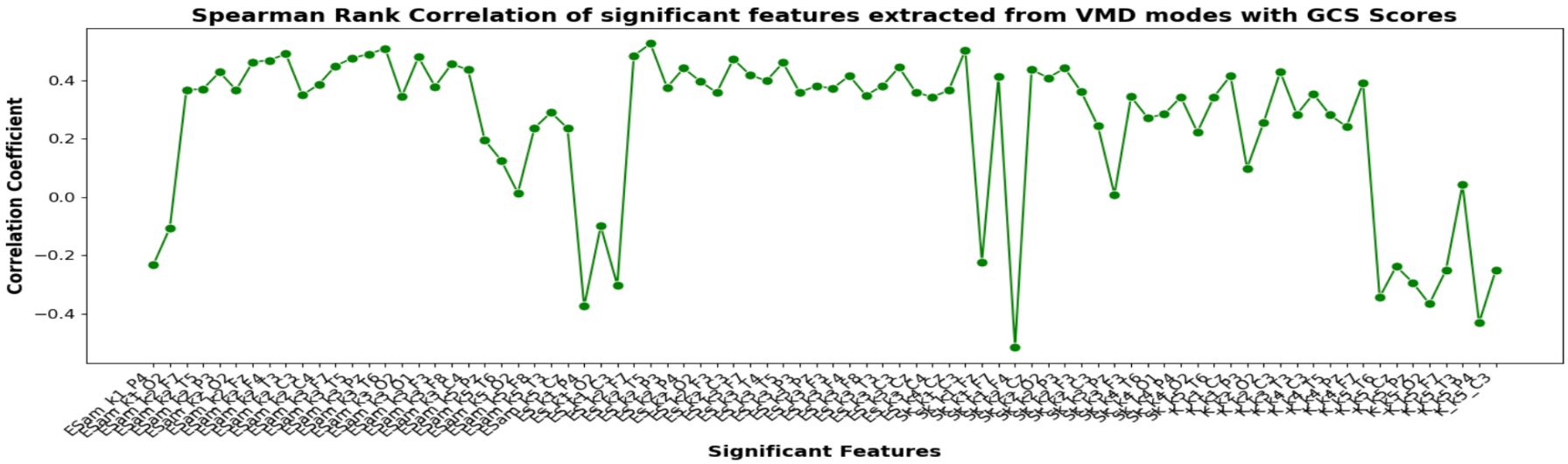



**Figure 4.11:** Spearman rank correlation coefficients (y-axis) between the GCS scores and selected features extracted from each of the five VMD modes

The results obtained for binary and multiclass classification using the reduced set of chosen features are shown for all the approaches in Tables 4.8 and 4.9, respectively. The results showed that the performance of the classifiers increased when only significant features were considered for classification. The utilization of significant features obtained using Kruskal-Wallis test in classification resulted in enhanced accuracy for all the approaches. The performance of the EBT classifier in the VMD-based approach saw an improvement from 83.3% to 86.7% in effectively distinguishing between UWS and MCS. For multiclass classification using EBT in VMD based approach, the performance increased from 76% to 80.5% with precision and recall of 81.1% and 81.7%. The F1 score improved from 78.6% to 81.4%.

Given the complexity of EEG signals, absolute correlation values exceeding 0.2 suggest that the corresponding EEG-derived features could have the potential to reflect improvements in patients' clinical conditions. Beyond conventional therapist assessments, these features warrant investigation as potential objective indicators of clinical status and may serve to evaluate the effectiveness of specific treatment strategies. If validated, such EEG-based metrics could constitute valuable tools for the objective clinical evaluation of patients with disorders of consciousness.

### 4.6.2 Improvement in the Performance of the Classifiers Due to Feature Selection

Table 4.10 illustrates the performance improvement achieved by excluding VMD-based features deemed non-significant according to the Kruskal-Wallis test. It demonstrates that feature selection leads to increased classification accuracy across all classifiers, with the exception of DT in the multiclass classification scenario and SVM in the binary classification scenario. By taking the performance of the EBT classifier beyond 80% in multiclass classification, this establishes the importance of feature selection in such machine learning problems. The inclusion of features that are not statistically significant confuses the classifier during training, leading to

**Table 4.8:** Binary (UWS vs MCS) classification results (in %) using 10-fold cross-validation with significant features extracted using Kruskal-Wallis test for the 3 different approaches. Number of patients: 30

| **Binary classification** | | | | | |
|---|---|---|---|---|---|
| | **Classifier** | **Accuracy** | **Precision** | **Recall** | **F1 score** |
| Approach 1: Features derived from the raw EEG | KNN | 56.7 | 52.5 | 57.5 | 54.9 |
| | Li.SVM | 50 | 49.2 | 52.5 | 50.8 |
| | DT | 43.3 | 40.8 | 47.5 | 43.9 |
| | EBT | 53.3 | 50 | 60 | 54.5 |
| Approach 2: Features derived from the frequency bands | KNN | 70 | 67.5 | 67.5 | 67.5 |
| | Li.SVM | 46.7 | 35.8 | 47.5 | 40.8 |
| | DT | 56.7 | 39.2 | 47.5 | 43 |
| | EBT | 76.7 | 70.8 | 77.5 | 74 |
| Approach 3: Features derived from the **VMD modes** | KNN | 73.3 | 60.8 | 72.5 | 66.1 |
| | Li.SVM | 80 | 68.3 | 75 | 71.5 |
| | DT | 66.7 | 63.3 | 65 | 64.1 |
| | **EBT** | **86.7** | **83.3** | **85** | **84.1** |

lower recognition performance. Thus, eliminating such features has the potential to improve the model's ability to generalize and perform equally well on new, unseen data.

### 4.6.3 Performance Comparison Between the Three Approaches Post Feature Selection

Table 4.11 compares the performance of the three approaches after feature selection based on the significance scores given by the Kruskal-Wallis test. The features derived from the VMD-based approach yield higher classification accuracy than those obtained from the frequency band-based and raw EEG approaches, regardless of the classifier used. Even when the optimal subset of features is selected for each method, VMD-derived features demonstrated a relative performance improvement of 50.5% and 61% over frequency-based and raw EEG-based features, respectively,

**Table 4.9:** Multi-class (Coma vs UWS vs MCS) classification results (in %) using 10-fold cross-validation with significant features extracted using Kruskal-Wallis test for the 3 different approaches. Number of patients: 45

| **Multiclass classification** | | | | | |
|---|---|---|---|---|---|
| | **Classifier** | **Accuracy** | **Precision** | **Recall** | **F1 score** |
| Approach 1: Features derived from the raw EEG | KNN | 44 | 37.2 | 45 | 40.7 |
| | Li.SVM | 32.5 | 21.6 | 36.7 | 27.2 |
| | DT | 55 | 50.3 | 56.7 | 53.3 |
| | EBT | 50 | 45.3 | 51.7 | 48.3 |
| Approach 2: Features derived from the frequency bands | KNN | 42 | 33.9 | 45 | 38.7 |
| | Li.SVM | 49 | 38 | 53.3 | 44.4 |
| | DT | 52 | 42.8 | 55 | 48.1 |
| | EBT | 53.5 | 44.7 | 56.7 | 50 |
| Approach 3: Features derived from the **VMD modes** | KNN | 61 | 53.1 | 58.3 | 55.6 |
| | Li.SVM | 70.5 | 67.8 | 75 | 71.2 |
| | DT | 57.5 | 50 | 58.3 | 53.8 |
| | **EBT** | **80.5** | **81.1** | **81.7** | **81.4** |

in the multiclass classification scenario. In the binary classification scenario, the corresponding improvements are 13

**Table 4.10:** Improvement in accuracy (%) of different classifiers due to the selection of features shown to be statistically significant by Kruskal-Wallis test: VMD-derived features. (IA: absolute improvement in accuracy)

| | Multiclass classification | | | Binary classification | | |
|---|---|---|---|---|---|---|
| Classifier | All features | Sel. features | IA | All features | Sel. features | IA |
| KNN | 46.0 | 61.0 | 15.0 | 60.0 | 73.3 | 13.3 |
| Li.SVM | 57.0 | 70.5 | 13.5 | 80.0 | 80.0 | 0 |
| DT | 73.5 | 57.5 | -16.0 | 63.3 | 66.7 | 3.4 |
| EBT | 76.0 | 80.5 | 4.5 | 83.3 | 86.7 | 3.4 |

The present study explored the effectiveness of variational mode decomposition as a preprocessing technique for resting-state EEG signals in patients with disorders of consciousness. The primary objective was to assess the viability of applying machine learning models for differentiating DOC patient groups, thereby providing

**Table 4.11:** Comparison of percentage increase in accuracy of VMD over frequency band-derived (FB) and raw EEG-derived (EEG) features after feature selection based on Kruskal Wallis test

| | Multiclass classification | | Binary classification | |
|---|---|---|---|---|
| Classifier | FB-features | Raw EEG-features | FB-features | Raw EEG-features |
| KNN | 45.2 | 38.6 | 4.7 | 29.3 |
| Li.SVM | 43.9 | 116.9 | 71.3 | 60 |
| DT | 10.6 | 4.5 | 17.6 | 54 |
| EBT | 50.5 | 61 | 13 | 62.7 |

an alternative to conventional behavior-based assessment methods, which are often subjective and prone to misclassification.

The research also sought to compare the efficacy of features derived from VMD with those extracted from traditional EEG frequency bands. The findings indicate that features extracted through VMD consistently outperformed those obtained from frequency band decomposition and raw EEG signals. The comparative improvements in classification accuracy are detailed in Tables 4.7 and 4.11, which present the percentage increases in classification accuracy using VMD-derived features. These comparisons were made both for the full set of extracted features and for statistically significant features identified through the Kruskal–Wallis test followed by Dunn's post-hoc analysis.

Table 6.12 provides a comparative overview of the proposed method with previously published approaches in EEG-based classification of DOC patients. The table summarizes key aspects such as feature types, classification models, and achieved accuracies across several studies by Höller et al., [36], Siit et al. [33], Engemann et al [34], Chennu et al. [17], and Di Gregorio et al., [31]. In work presented by Höller et al., [36] employed directed transfer function features in conjunction with Li.SVM classifiers, achieving classification accuracies of 82% for distinguishing between MCS and UWS patients, and 92% for healthy controls (HC) versus MCS. Similarly, Siit et al. in their study [33] integrated features such

as wSMI, PE, spectral entropy, and KCC with SVM classifiers, attaining a high accuracy of 89% for MCS versus UWS classification. This study also provided an extensive characterization of EEG signal features. In contrast, Engemann et al. in their work [34] combined normalized alpha-band power, PE, wSMI, and complexity measures for classification using SVMs, though specific performance metrics were not explicitly reported.

These comparative results underscore the potential of the VMD-based preprocessing framework in enhancing the performance of automated EEG-based classification systems for DOC assessment. The findings suggest that VMD enables the extraction of more discriminative and informative signal components, thereby contributing to improved classification performance across both binary and multi-class DOC differentiation tasks.

In the study by Chennu et al. [17], a combination of spectral power, debiased weighted phase lag index, and graph theoretic metrics was utilized in conjunction with a radial basis support vector machine (Rb.SVM) classifier to differentiate among various states of consciousness namely, UWS, minimally conscious state minus (MCS−), minimally conscious state plus (MCS+), and healthy controls achieving an overall classification accuracy of 79%. Similarly, the work presented by Di Gregorio et al., in [31] employed features such as mutual information (MI), dominant frequency, and partial coherence alongside LDA classifiers to distinguish between TBI and non-TBI patients. While this study reported varying accuracies across different TBI subgroups, it did not specifically address the classification of DOC categories. Notably, both studies focused exclusively on binary classification tasks and did not incorporate coma patients into their analysis.

**Table 4.12:** Comparison of performance of proposed method with existing work

| **Author** | **Features** | **No.of Subjects** | **Classifier** | **Accuracy (%)** | **Precision (%)** |
|---|---|---|---|---|---|
| Holler et al. (2014) | directed transfer function | MCS 22, UWS 27, Controls 23 | Li.SVM | MCS vs. UWS: 82 | - |
| | generalized partial directed coherence | | | HC vs. MCS: 92 | - |
| | PCoh | | | HC vs. UWS: 92.6 | - |
| Sitt et al. (2014) | wSMI, PE, SpecE, KCC | MCS 68, UWS 75 | SVM | 89 | - |
| Engemann et al. (2015) | normalized alpha-band power, PE, wSMI, complexity measure | MCS 69, UWS 76 | SVM | - | - |
| Chennu et al. (2017) | spectral power, dwPLI, graph-theoretic metrics | UWS 23, MCS- 17, MCS + 49, controls 26 | Rb.SVM | 79 | - |
| Di Gregorio et al. (2022) | dominant frequency, PCoh | DOC 33 | LDA | TBI: 80 | 77.7 |
| | MI, PCoh | | | Non-TBI: 83.3 | 85.7 |
| **Proposed Method** | VMD mode based features: kurtosis, skewness, spectral | Coma 15, UWS 15, MCS 15 | Ensemble bagged tree | UWS vs. MCS: **86.7** | **83.3** |
| (features selected by Kruskal-Wallis test) | entropy, sample entropy | | | Coma vs. UWS vs. MCS: **80.5** | **81.1** |

The proposed method, in contrast, introduces a novel application of VMD-derived features in combination with an EBT classifier to distinguish among coma, UWS, and MCS patients. This approach yielded a promising accuracy of 86.7% in the binary classification between UWS and MCS groups, and 80.5% accuracy in the multi-class classification task encompassing coma, UWS, and MCS. These results

underscore the potential of VMD-based features, in conjunction with ensemble learning techniques, to enhance classification performance in the context of DOC especially in the more complex, multi-class setting that includes coma patients, which previous studies have largely overlooked.

One of the key advantages of VMD-based feature extraction lies in its ability to capture transient and non-stationary dynamics in EEG signals phenomena commonly observed in DOC patients, who may exhibit fluctuating levels of consciousness or intermittent neural responsiveness. Traditional frequency band analysis, while useful, may lack the sensitivity to detect such temporal variations. VMD, by contrast, decomposes EEG signals into a predefined number of modes that can adapt to local signal characteristics, thereby isolating subtle shifts in neural activity that may correspond to changes in patient state.

VMD facilitates enhanced spatial localization of neural activity by producing modes that are more spatially independent, which can assist in identifying abnormal brain regions and inform interventions such as neuromodulation or neurofeedback. Another notable benefit is its flexibility across EEG acquisition setups and patient-specific variability. Unlike conventional frequency band decomposition, VMD does not rely on rigid, pre-established frequency bands, making it more adaptable to diverse EEG protocols and individual patient neural dynamics. This adaptability increases its applicability in both research and clinical settings, enhancing its value as a diagnostic and prognostic tool.

Although VMD has demonstrated superior performance across various domains involving complex EEG signal analysis including non-clinical applications it has not, to the best of our knowledge, been previously applied in the classification of DOC patient groups. This study, therefore, addresses a critical gap in the existing literature and provides empirical evidence supporting the utility of VMD in the analysis of resting-state EEG for DOC assessment.

It is important to emphasize that VMD is not intended to replace traditional

EEG frequency band analysis, but rather to complement it. While frequency bands offer a coarse-grained view of consciousness and cognitive functioning, VMD enables a more nuanced, dynamic analysis of the signal. A hybrid approach that integrates both methodologies may yield the most comprehensive understanding of brain activity in DOC patients, facilitating improved diagnostic accuracy and potentially informing patient management strategies.

## 4.7 Summary of the Chapter

This chapter addressed the challenges of classifying patients with disorders of consciousness using resting-state EEG signals, which are complex and non-stationary in nature. It proposed variational mode decomposition as a preprocessing technique to enhance feature extraction and classification accuracy by decomposing EEG signals into distinct modes and extracting features such as sample entropy, spectral entropy, kurtosis, and skewness. The classification framework evaluated the performance of VMD-derived features against those from traditional frequency band-based methods and raw EEG data in both binary (UWS vs. MCS) and multi-class (coma vs. UWS vs. MCS) scenarios. Statistically significant features, identified using the Kruskal-Wallis test ($p < 0.05$), were utilized for final classification experiments. The results demonstrated that the VMD-based approach performed better than the alternative methods, with the ensemble bagged tree classifier achieving higher accuracy in both multi-class and binary classification. These findings underscored the effectiveness of VMD in improving classification accuracy and highlighted its potential for enhancing the diagnosis and management of DOC patients, reinforcing the value of integrating advanced signal processing techniques with machine learning in clinical applications.

**The following research papers were published based on this chapter:**

1. Sreelakshmi Raveendran, et al., "Variational mode decomposition-based EEG analysis for the classification of disorders of consciousnes" in *Frontiers in Neuroscience*, 2024, 18,

1340528.

2. Sreelakshmi Raveendran, et al., "Scalp EEG-based Classification of Disorder of Consciousness States using Machine Learning Techniques" in *IEEE, 11th International Symposium on Electronic Systems Devices and Computing (ESDC 2023)*, IEEE, Sri City, India, 2023.

# Chapter 5

# Regression Models to Predict Consciousness Level

This chapter presents an in-depth investigation about the predictive capacity of conventional regression models in estimating the level of consciousness in patients with disorders of consciousness, based on features extracted from resting-state EEG signals. By leveraging quantitative EEG-derived features the study explores how well these models can approximate clinically relevant assessments of patient awareness. Along with evaluating the performance of classic regression algorithms, this chapter introduces a novel predictive framework: an ensemble of ensemble regressor model (EoERM) by integrating XGBoost, random forest, and support vector regressor using stacking method to predict the consciousness states of DOC patients. To validate the clinical utility of the predicted values, the estimated consciousness scores generated by the regression models are normalized and compared to the CRS-R scores for each patient. The clinical relevance and applicability is assessed by aligning the predicted consciousness level to clinical score. The agreement is measured using cohen's kappa coefficient.

## 5.1 Feature Extraction

Symbolic transfer entropy (STE), along with absolute and relative power were the features analyzed [23]. These features were extracted from 17 EEG channels across five frequency bands (delta, theta, alpha, beta, and gamma) and aggregated based on anatomically defined brain regions. The regional aggregation enabled a spatially informed analysis of EEG dynamics across distinct frequency bands, facilitating interpretation in terms of cortical information processing and inter-regional connectivity.

Among the features considered, STE represents a powerful information theoretic metric used to assess the directional transfer of information between two time series signals. STE is a symbolic extension of the traditional transfer entropy (TE), originally introduced by Schreiber [76, 77], and is applied for analyzing nonlinear, non-stationary, and noisy signals, like EEG.

Transfer entropy quantifies the information flow from a source process $X$ to a target process $Y$ by comparing the predictability of $Y$'s future values with and without the historical knowledge of $X$. It is formally defined as:

$$TE(X \to Y) = H(Y_t|Y_{t-1}, Y_{t-2}, \ldots) - H(Y_t|Y_{t-1}, Y_{t-2}, \ldots, X_{t-1}, X_{t-2}, \ldots) \quad (5.1)$$

Here, $H(\cdot)$ denotes the Shannon entropy, and the equation captures how much additional predictive power is gained for the future state $Y_t$ by incorporating the past values of process $X$, over and above using only the past of $Y$. However, calculating TE directly from continuous valued EEG signals can be computationally intensive and sensitive to noise. To address this, symbolic transfer entropy transforms the continuous EEG time series into symbolic representations via discretization techniques such as binning or ordinal pattern encoding. This symbolic transformation reduces sensitivity to outliers and improves robustness, particularly in short or

noisy EEG segments. Once symbolic sequences are constructed for both processes $X$ and $Y$, entropy of the symbolic sequences is computed to assess the uncertainty or information content. The marginal entropy of the symbolic target sequence $S_Y$ is given by:

$$H(S_Y) = -\sum_{y \in S_Y} P(y) \log P(y) \tag{5.2}$$

To evaluate the influence of $X$ on $Y$, the conditional entropy of the symbolic sequence $S_Y$, conditioned on its own past and the past of $X$, is computed as:

$$H(S_Y | S_{Y\ \text{past}}, S_{X\ \text{past}}) = -\sum_{y, y_{past}, x_{past}} P(y, y_{past}, x_{past}) \log P(y | y_{past}, x_{past}) \tag{5.3}$$

The final expression for symbolic transfer entropy from $X$ to $Y$ is then given by:

$$STE(X \to Y) = H(S_Y | S_{Y\ \text{past}}) - H(S_Y | S_{Y\ \text{past}}, S_{X\ \text{past}}) \tag{5.4}$$

This formulation quantifies the reduction in uncertainty about the future state of $Y$ when incorporating the past of $X$, beyond what is already explained by the past of $Y$ alone. A higher STE value indicates a stronger directional influence from $X$ to $Y$, interpreted as an information transfer [76].

In the context of this research, STE was computed between pairs of EEG channels across multiple frequency bands, allowing for the construction of functional connectivity profiles within and across brain regions. These profiles are especially relevant in DOC research, where altered directional connectivity patterns can indicate impaired consciousness related network dynamics. By aggregating these measures regionally, the study aimed to capture patterns of directional information flow to distinguish between coma, UWS, and MCS.

## 5.2 Prediction Methods and Performance Metrics

### 5.2.1 Prediction Methods

The regression models used for predicting the consciousness levels are linear regression, elastic net, extreme gradient boost, SVR, and random forest [69].

The SVR model was implemented with the C parameter set at 1.0, balancing low training error with model complexity, thereby influencing generalization. The model employed a radial basis function kernel for effective non-linear relationship handling. The epsilon parameter is established at 0.1, defining the width of the epsilon insensitive zone to ignore minor errors.

The random forest model is configured with the bootstrap parameter set to true, allowing for bootstrap sampling of data. The criterion is defined as squared error, optimizing the split quality by minimizing squared differences between predictions and actual values. The model is implemented with 100 trees and is initialized with a random state of 42.

Elastic net regression utilizes ridge and lasso regression properties, making it useful for dealing with multicollinear features and performing variable selection [78]. The alpha parameter is configured at 0.1, regulating the intensity of regularization. The l1 ratio is designated as 0.1. The maximum number of iterations during optimization is determined to be 1000. The random state is set to 42 to ensure reproducibility, and the selection method is established as cyclic, signifying that coordinate descent will be executed cyclically.

The linear regression model is configured with key parameters that affect its training behavior. The fit intercept parameter is set to true to include an intercept term, ensuring accurate fitting of data that does not pass through the origin. The copy X parameter is also true, allowing the model to make a copy of the input data to avoid altering the original dataset.

Extreme gradient boost which is an ensemble learning technique has a series of decision trees sequentially connected, where each tree corrects the errors of its predecessor. The extreme gradient boost parameters include decision trees with a depth of 3 layers, 100 estimators, and a learning rate of 0.1.

### 5.2.2 Performance Metrics

The performance of each regression model was evaluated using three statistical metrics namely mean squared error (MSE), mean absolute error (MAE), and coefficient of determination ($R^2$) [79]. These metrics provide complementary insights into the predictive accuracy and generalization capability of the models [80].

- **Mean Squared Error:** MSE is one of the most commonly employed loss functions in regression analysis and is also referred to as the L2 loss. It quantifies the average squared difference between the observed actual values and the values predicted by the model. The squaring operation penalizes larger errors more heavily than smaller ones, making MSE sensitive to outliers. A lower MSE value, ideally approaching zero, indicates a higher degree of accuracy and a better model fit. The MSE is mathematically expressed as:

$$\text{MSE} = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 \tag{5.5}$$

  where $y_i$ represents the actual target value, $\hat{y}_i$ denotes the predicted value, and $m$ is the total number of observations [79].

- **Mean Absolute Error:** MAE, also known as the L1 loss, measures the average absolute difference between the actual and predicted values, without considering the direction of the error. Unlike MSE, it treats all errors linearly, giving equal weight to all deviations, regardless of magnitude. This makes MAE more robust to outliers [79]. It is useful when the interpretability of

the average error magnitude is desired. MAE is defined as:

$$\text{MAE} = \frac{1}{m}\sum_{i=1}^{m}|y_i - \hat{y}_i| \tag{5.6}$$

As with MSE, smaller MAE values are indicative of more accurate predictions.

- **Coefficient of Determination:** The $R^2$ metric, also known as the coefficient of determination, quantifies the proportion of variance in the dependent variable that is predictable from the independent variables. It serves as a measure of the explanatory power of the model. An $R^2$ value of 1 implies that the regression model perfectly predicts the observed data, whereas a value of 0 indicates that the model does no better than simply predicting the mean of the target variable. Negative values may occur when the model performs worse than a horizontal mean line [79]. The $R^2$ score is computed as:

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{m}(y_i - \overline{y})^2} \tag{5.7}$$

where $\overline{y}$ is the mean of the actual values. The range of $R^2$ extends from $-\infty$ to 1, with values closer to 1 denoting better model performance.

Together, these evaluation metrics provide a comprehensive assessment of model effectiveness by capturing different aspects of prediction error magnitude, variance explained, and sensitivity to outliers.

## 5.3 Ensemble of Ensemble Regressor Model: Eo-ERM

An ensemble of ensemble regressor model was developed to predict the consciousness levels of DOC patients, utilizing features extracted from rs-EEG data in conjunction with clinical behavioral assessment scores recorded at the time of admission. This modeling approach integrated multiple machine learning algorithms through a

stacking ensemble framework, which has been demonstrated to enhance predictive performance by leveraging the complementary strengths of diverse base learners [81].

The stacking architecture comprised three base-level regressors XGBoost, random forest regressor, and SVR; each independently trained on the same input feature set. The predictions generated by these base models were then passed to a linear regression model functioning as a meta-learner. The meta-learner synthesized the predictions to generate the final output, effectively learning how to optimally combine the individual models' outputs to improve overall performance. The choice of the three regression models employed in constructing EoERM was determined based on the outcomes detailed in Table 5.2, as well as the combination of regression models that demonstrated better performance with the selected set of EEG-derived features.

This hierarchical ensemble design leveraged the unique strengths of each individual regressor to enhance overall modeling performance. XGBoost and random forest are both tree-based ensemble methods known for their robustness and effectiveness in capturing complex, non-linear relationships and feature interactions within high-dimensional data. SVR, on the other hand offers strong generalization properties, particularly in scenarios characterized by data variations and smaller sample sizes. The integration of these models through stacking enabled the ensemble to achieve a balance between robustness, flexibility, and precision [81]. The parameter settings of the models integrated into EoERM are consistent with those employed in the individual performance evaluations. The hyperparameter configurations for each model were selected in a data-driven manner, through iterative experimentation and cross-validation, to identify the optimal settings that yielded the highest predictive accuracy. This methodical approach ensured that EoERM model not only capitalized on the complementary capabilities of the individual regressors but also adapted effectively to the nuances of EEG-derived features and the clinical profiles of DOC patients. The graphical depiction of the methodology adopted is shown in

Figure 5.1.

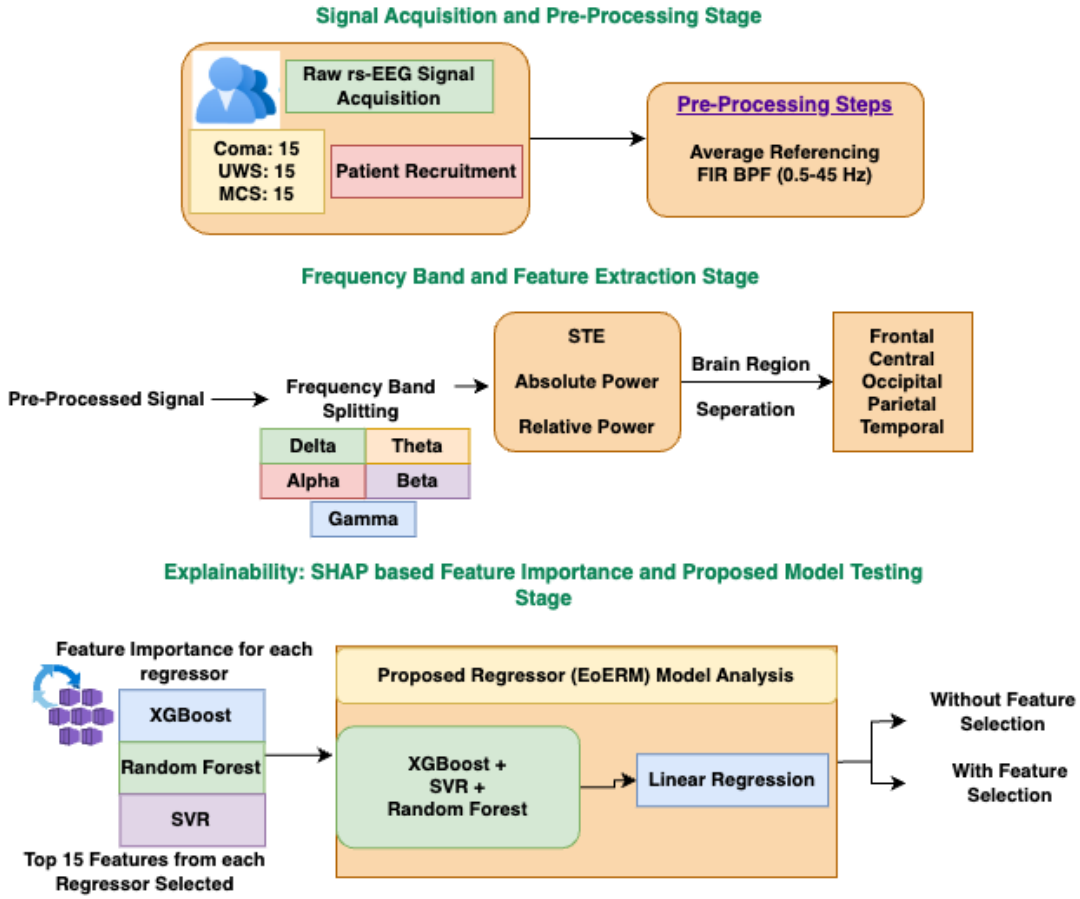



**Figure 5.1:** Graphical depiction of the proposed methodology

## 5.4 Explainability of Features and Feature Selection using SHAP

To enhance the interpretability of the model and elucidate how input features influence the final predictions, an explainable artificial intelligence (XAI) framework was employed. XAI methodologies are designed to provide transparent and comprehensible insights into the internal decision-making processes of machine learning models, thereby enabling a deeper understanding of how specific inputs contribute to the model's outputs. In the context of predictive modeling, the utility of XAI lies in its ability to identify and quantify the influence of individual features

or variables on the model's predictions.

SHAP method was utilized as the primary XAI technique. SHAP is grounded in cooperative game theory and assigns each input feature a SHAP value, which represents the marginal contribution of that feature to a specific prediction. These values offer a consistent and locally accurate measure of feature importance, enabling the decomposition of a model's output into additive contributions from each input variable.

Using SHAP analysis, the top 15 features with the highest importance scores were extracted from each of the base regression models employed in the ensemble. These features were then aggregated to construct a comprehensive feature importance vector, which served as the basis for subsequent analysis, including feature selection and the development of the consciousness level computation framework. After eliminating redundancies, a total of 32 distinct and highly influential features were identified as critical to the model's predictive performance.

Table 5.1 presents the list of selected features along with their respective importance measures as determined by the SHAP analysis. These features reflect diverse neurophysiological patterns derived from resting-state EEG data and were instrumental in enhancing both the model's accuracy and interpretability.

## 5.5 Consciousness Level Estimation Based on Clinical Assessment and EEG Features

Assessment of consciousness level of each subject in the clinical environment is determined using CRS-R scores during various stages like at the time of admission of the patient, after 28 days of the being in DOC stae, after 3 months and 6 months. CRS-R scores range from 0 to 23, with higher values indicating higher levels of conscious awareness and cognitive function [13]. CRS-R scores above 10 typically indicate the presence of MCS or emergence from MCS. Patients diagnosed with coma

**Table 5.1:** Combined feature importance from each of the 3 regressors calculated using SHAP

| **Features** | **Importance** | **Features** | **Importance** |
|---|---|---|---|
| alpha_parietal_ab | 0.10306 | delta_central_ab | 0.022832 |
| theta_central_ste | 0.09434 | delta_parietal_ab | 0.022693 |
| theta_central_rp | 0.066908 | alpha_occipital_rp | 0.02249 |
| beta_central_rp | 0.064735 | delta_occipital_ab | 0.020982 |
| beta_frontal_rp | 0.06282 | delta_central_rp | 0.020936 |
| theta_occipital_ste | 0.051224 | beta_temporal_ab | 0.020011 |
| beta_occipital_rp | 0.038823 | delta_frontal_rp | 0.019543 |
| beta_temporal_rp | 0.034787 | theta_central_ab | 0.019412 |
| delta_central_ste | 0.030414 | alpha_occipital_ab | 0.018799 |
| alpha_parietal_rp | 0.029839 | alpha_frontal_rp | 0.01871 |
| theta_parietal_ab | 0.029598 | gamma_parietal_ste | 0.017557 |
| alpha_frontal_ab | 0.026626 | alpha_frontal_ste | 0.014265 |
| delta_midline_ste | 0.026269 | delta_frontal_ab | 0.01405 |
| gamma_occipital_ste | 0.024016 | beta_temporal_ste | 0.009732 |
| beta_parietal_rp | 0.02399 | beta_midline_ste | 0.009345 |
| delta_temporal_ste | 0.023649 | alpha_central_rp | 0.008528 |

generally exhibit profound impairments in both awareness and wakefulness, often reflected by CRS-R scores ranging between 0 and 3. Patients in the UWS category display some degree of wakefulness but lack awareness of self or environment, and their scores usually fall between those observed in coma and MCS groups [13, 82].

However, the enrolled patient cohort exhibited a degree of overlap in CRS-R scores between UWS and MCS categories, complicating the rigid separation of these groups based on behavioral assessment method alone. This phenomenon reflects the clinical complexity and continuum of consciousness often observed in

DOC populations. To address this challenge and explore the potential of EEG-derived features in enhancing consciousness assessment and predicting consciousness level, a data-driven approach for predicting consciousness levels using rs-EEG in conjunction with machine learning techniques were proposed.

A customized equation to compute an individualized consciousness level (denoted as $L_C$) for each patient was developed. This equation was constructed using the most informative EEG features, identified through SHAP analysis, which ranks features based on their contributions to model predictions. The top-ranked EEG features from Table 5.1 mainly included STE and relative power values from multiple brain regions including parietal, temporal, frontal, and central lobes. The weights of the top featured were calculated using linear regression model. The resulting regression coefficients ($\beta_i$) quantified the relative contribution of each feature ($X_i$) to the overall consciousness level. The intercept term ($\beta_0$) and the coefficients were learned from the data using ordinary least squares fitting, optimizing the fit between EEG-derived features and the ground truth clinical assessments. The resulting consciousness level score $L_C$ is computed using the following linear model:

$$L_C = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_n X_n \tag{5.8}$$

Here, $\beta_0$ represents the model intercept, which was determined to be 3.31 based on training, while $\beta_1$ through $\beta_n$ are the respective coefficients for the top $n$ features, and $X_1$ through $X_n$ denote the feature values for a given patient. The value of $n$ was selected as 15. Following computation, the raw consciousness level scores were normalized to the standard CRS-R scale (0 to 23) to facilitate clinical interpretability and comparison. This normalization step ensures that the model-derived scores remain within the clinically meaningful bounds defined by the CRS-R, thus aligning the EEG-based measure with established behavioral assessment standards.

## 5.6 Results

### 5.6.1 Performance of Basic Regressor Models to Predict Consciousness Level

The performance of the regression models developed was systematically evaluated using features extracted from rs-EEG signals, including absolute power, relative power, and symbolic transfer entropy. These features were selected for their ability to reflect the underlying neural mechanisms associated with varying levels of consciousness in DOC patients. The models were trained and tested on this feature set to predict individualized consciousness levels.

Table 5.2 presents a comparative summary of the performance metrics for each of the regression models investigated in this study.

**Table 5.2:** Comparison of different regression models based on performance metrics

| **Regressor** | **MSE** | **MAE** | $R^2$ |
|---|---|---|---|
| SVR | 0.325 | 0.477 | 0.576 |
| Random Forest | 0.232 | 0.406 | 0.697 |
| Elastic Net | 0.184 | 0.387 | 0.759 |
| Linear Regression | 0.143 | 0.31 | 0.813 |
| **XGBoost** | **0.111** | **0.114** | **0.855** |

The performance analysis of the implemented regression models revealed varying degrees of predictive capability in estimating consciousness levels from EEG-derived features. Among the models evaluated, the SVR exhibited comparatively lower performance, characterized by elevated MSE and MAE values, and a relatively low $R^2$ of 0.576. Although the SVR was able to capture certain nonlinear relationships between EEG features and consciousness levels, its predictive accuracy was limited relative to the other regression models, including elastic net, extreme gradient boosting, random forest, and linear regression.

Random forest model also underperformed in comparison to the other regressors,

achieving an $R^2$ value of 0.697 and an MAE of 0.406. While the ensemble nature of the random forest allows for the modeling of complex data interactions and reduction of variance through bootstrap aggregation, its effectiveness in this context was suboptimal, possibly due to high variance in the feature set or overfitting tendencies in high-dimensional input spaces.

Elastic net regression model, an integration of lasso and ridge regularization methods demonstrated moderate predictive accuracy, achieving an $R^2$ value of 0.759. This result indicates that approximately 75.9% of the variance in the consciousness level can be accounted for by the EEG features used. The model's corresponding MSE and MAE values were within acceptable ranges, signifying reasonable predictive capacity. However, it also highlighted the typical trade-off between model simplicity and predictive strength, with elastic net offering a balance between interpretability and performance.

Linear regression model yielded improved results in comparison to SVR, random forest, and elastic net regressors. It achieved an $R^2$ value of 0.813 and a lower MAE of 0.31, reflecting better overall model fit and reduced prediction error. Despite its simplicity and assumption of linearity, the model was able to capture essential trends in the data, indicating the presence of strong linear relationships between selected EEG features and consciousness levels.

Among all models evaluated, XGBoost regressor demonstrated superior performance. It achieved the highest $R^2$ value of 0.855, thereby explaining 85.5% of the total variance in consciousness levels. The model also reported the lowest MAE and MSE values, at 0.114 and 0.111 respectively, emphasizing its high precision and low error rate. These results underscore the robustness of the XGBoost regressor in modeling complex, nonlinear dependencies between multivariate EEG features and behavioral assessments. Its ensemble boosting architecture enables it to iteratively correct prediction errors and focus on difficult-to-model patterns, making it highly suitable for this application. The effectiveness of the XGBoost model highlights

its potential utility in clinical settings for diagnostic and prognostic applications related to disorders of consciousness.

### 5.6.2 Results of the Proposed Model without Feature Selection

In the baseline configuration, EoERM was trained and evaluated using the complete set of EEG features without applying any feature selection technique. Under this setting, the model demonstrated a better capacity to predict consciousness levels with notable precision. The calculated MSE for this model configuration was 0.090, indicating a relatively low average of squared prediction errors. The MAE was computed as 0.235, signifying that the average magnitude of prediction error across all samples was approximately 0.235 units. The $R^2$ value obtained for this model was 0.88, which implies that 88% of the variability in the observed consciousness levels could be attributed to the features included in the model. This coefficient of determination reflects the model's effectiveness in leveraging the information embedded in EEG signal features, even without the refinement introduced by feature selection.

These results establish the baseline performance of EoERM and demonstrate its potential for the prediction of consciousness levels. However, the inclusion of all features, regardless of their relevance, contributes to noise and redundancy in the model, thereby limiting its generalizability. Evaluations incorporating feature selection were intended to address these limitations and further optimize the model's performance.

### 5.6.3 Results of the Proposed Model with Feature Selection

To address the aforementioned limitations and refine the model's predictive accuracy, XAI methods were employed for feature selection. In particular, SHAP method was used to identify and rank the most influential features across the trained regressors.

From each regression model, the top 15 features with the highest SHAP values were extracted and integrated to construct an optimized feature vector for the proposed ensemble model. Figures 5.2, 5.3, and 5.4 present the SHAP summary plots for each base model used in the ensemble, offering insight into the relative importance of EEG features in determining consciousness levels.

The performance of the proposed model significantly improved following the application of this SHAP-based feature selection. MSE was reduced to 0.0356, a substantial improvement over the baseline value of 0.0909. This indicates a marked reduction in the average squared prediction error, suggesting that the model was now better equipped to generalize to unseen data. MAE decreased to 0.151, in contrast to the baseline MAE of 0.235, implying that the model's average deviation from the true values had been reduced by approximately 36%.

The $R^2$ value increased to 0.95, indicating that 95% of the variance in the predicted consciousness levels could be explained by the selected EEG features. This improvement underscores the efficacy of the feature selection approach in isolating the most relevant neurophysiological markers and eliminating noise from the model's input space.

The improvement observed across all three evaluation metrics MSE, MAE, and $R^2$ demonstrated the importance of feature selection. By filtering out redundant or weakly informative features, the model was better able to identify meaningful relationships between EEG patterns and behavioral consciousness assessments. The improved precision, as evidenced by the lower error rates, coupled with the increased explanatory power, as indicated by the elevated $R^2$, highlights the robustness of the refined model. This version offers a more reliable tool for clinical application, particularly in scenarios requiring accurate consciousness level estimation from non-invasive EEG data.

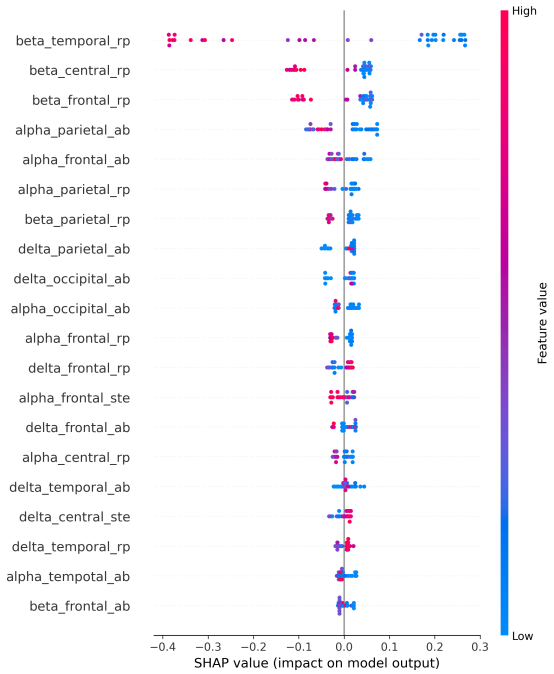



**Figure 5.2:** Summary plot obtained using SHAP method for random forest regressor model

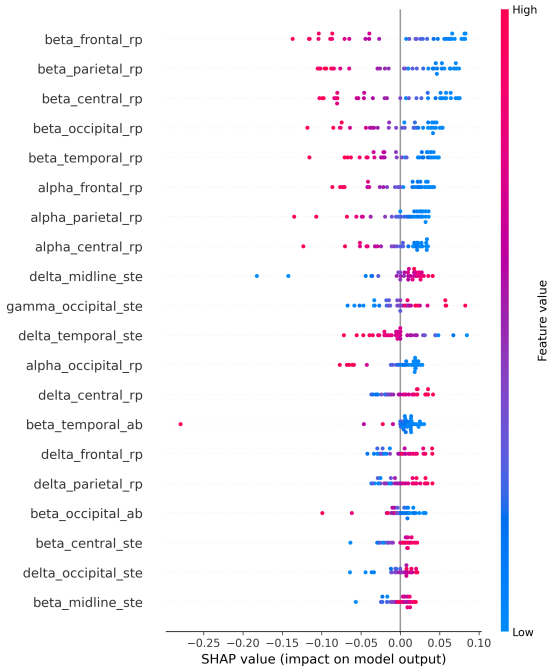



**Figure 5.3:** Summary plot obtained using SHAP method for SVR regressor model

### 5.6.4 Normalized Consciousness Level-Based Analysis

In a further evaluation, the proposed prediction model was tested on the consciousness level data of each subject normalized to match the standardized 0–23

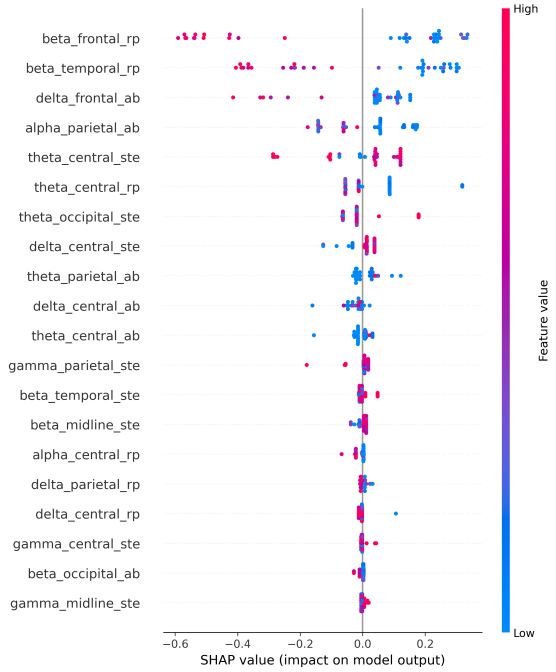



**Figure 5.4:** Summary plot obtained using SHAP method for XGBoost regressor model

scale of the CRS-R score. This approach provided a framework for direct clinical interpretability. However, the model exhibited higher prediction error metrics in this configuration. MSE increased to 4.561 from the baseline value of 0.0909, indicating higher average squared deviation between predicted and actual normalized consciousness levels. Similarly, the MAE rose to 1.69, suggesting that the model's predictions deviated by approximately 1.69 points on the CRS-R scale from the ground truth, on average.

Despite these higher error rates, the model achieved an $R^2$ value of 0.91, demonstrating it ability to capture 91% of the variance in the normalized consciousness levels. This finding implied that, while individual predictions could lack precision, the model retained the capacity to uncover and model the underlying trends in the data. Moreover, the model exhibited a 97% agreement with clinical CRS-R classifications, demonstrating strong consistency with clinical assessments. A Cohen's kappa coefficient of 0.97 indicates near-perfect agreement, and the bootstrapped p-value of 0.00001 further supports the statistical significance of this outcome.

The distribution of model-derived consciousness scores was visualized using a violin plot, constructed based on the normalized consciousness levels across three diagnostic groups: coma, UWS, and MCS. The plot, shown in Figure 5.5, offers a nuanced depiction of the density and distribution of consciousness scores within each group. The MCS group displayed a broader and higher distribution, indicating greater variability and elevated consciousness levels compared to the coma and UWS groups. Importantly, the predicted consciousness scores aligned with established CRS-R boundaries, supporting the validity of EEG-based assessments as complementary tools to traditional behavioral scales.

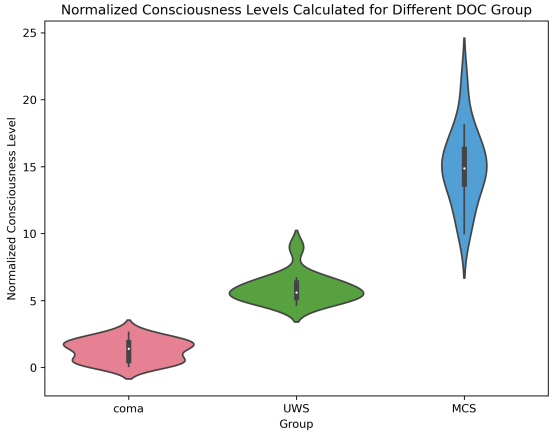



**Figure 5.5:** Normalised consciousness levels for the three groups of DOC patients.

A comparative overview of performance metrics for all three model configurations baseline (without feature selection), SHAP-based feature selection, and normalized analysis is provided in Table 5.3.

**Table 5.3:** Performance of the proposed regressor EoERM with and without feature selection task, as well as normalized consciousness level calculated using the custom equation.

| **Ensemble of Ensemble Regressor Model** | **MSE** | **MAE** | **$R^2$** |
|---|---|---|---|
| Proposed Model (without feature selection) | 0.0909 | 0.235 | 0.88 |
| Proposed Model (with feature selection) | 0.0356 | 0.151 | 0.95 |
| Normalized Consciousness Level based Analysis | 4.561 | 1.69 | 0.91 |

## 5.7 Summary of the Chapter

This chapter presented a methodologically robust and data-driven framework for estimating consciousness levels in DOC patients using rs-EEG data. By implementing advanced regression techniques and integrating XAI-based feature selection, the study achieved better predictive performance also offered interpretable models that can potentially support clinical decision-making. The comparative analysis showed that models enriched with selected EEG features performed better than full feature counterparts, both in terms of precision and generalization. The integration of normalized consciousness scores facilitated alignment with clinical scales, enabling potential translational applications in real-world diagnostic and prognostic scenarios. The proposed methodology emphasized the importance of combining explainability with predictive modeling in healthcare domains, particularly in sensitive and complex areas such as DOC assessment.

**The following research paper was published based on this chapter:**
Sreelakshmi Raveendran, et al., "Machine learning-based consciousness assessment in disorders of consciousness patients" in *International Conference on Brain-Computer Interface and Health Technologies (iCon-BCIHT 2024)*, IEEE, Thiruvananthapuram, India, 2024.

# Chapter 6

# Functional Connectivity Analysis and Deep Learning-Based Classification of DOC Patients

Understanding the neural connections and correlations among different DOC cohorts is essential for developing advanced diagnostic methods. This chapter is organized into two main sections. In the first section 6.3.1, various deep learning models are implemented to classify patients across different DOC groups. These models are trained using FC features namely PLI, wPLI, and AEC extracted from preprocessed EEG data. To capture dynamic neural activity, varying temporal window sizes are employed during feature extraction.

The second section 6.4.2 focuses on the analysis of functional brain networks. Functional connectivity measures such as dwPLI, coherence, and AEC are used to construct undirected weighted brain networks, where the connectivity metrics serve as edge weights. From these networks, six graph theoretical features are extracted to investigate differences in brain network organization across DOC cohorts and to provide insights into underlying neural mechanisms. To validate the clinical implication of the results, CCA is employed to quantify the relationship between EEG network features and GOSE score obtained at three months.

The following are the major contributions of this chapter:

- Evaluation of multiple ANN architectures (MLP, RNN, LSTM, GRU, CNN-LSTM) to identify effective model for DOC classification.
- Analysis of effect of varying temporal window lengths (16–20 seconds) on classification performance to determine the optimal resolution for neural dynamics.
- Assessment of confidence intervals to assess model robustness and reliability by quantifying uncertainty and consistency.
- Undirected weighted functional networks with AEC, dwPLI, and coherence generation for the analysis of neural interactions in DOC patients.
- Examination of relationship between EEG network features and GOSE scores at three months using canonical correlation analysis for better prognostic insights.

## 6.1 Feature Extraction

Functional connectivity refers to the statistical dependencies or correlations between spatially distinct brain regions, reflecting their interaction during rest or task performance [83].

Resting-state EEG-based FC analysis provides a non-invasive, high temporal resolution approach to evaluate brain connectivity in DOC patients. FC features extracted from EEG signals include phase locking value, coherence, phase lag index, weighted phase lag index, and debiased weighted phase lag index which assess synchronization and directional interactions between brain regions [11]. Graph theoretical analysis further elucidate network organization, highlighting disruptions in connectivity patterns characteristic of DOC. Studies indicate that DOC patients, particularly those in a vegetative state, exhibit decreased long range connectivity and increased local connectivity, particularly within the DMN and frontoparietal networks [17]. These connectivity alterations correlate with the

severity of consciousness impairment and provide prognostic insights, as preserved or enhanced FC is often associated with greater recovery potential.

### 6.1.1 Functional Connectivity Measures

#### Phase Lag Index

Phase lag index is a metric that measures functional connectivity by assessing the asymmetry of phase differences between two signals over time [84]. PLI focuses on the consistency of non-zero phase lags, making it robust against the effects of volume conduction and common source artifacts. This robustness is particularly beneficial in the study of brain connectivity, where the elimination of spurious correlations is essential for accurately identifying genuine neural interactions [84].

To compute PLI, the first step is to extract the instantaneous phase of a signal. Given a real valued time series $x(t)$, its analytic signal is obtained using Hilbert transform $z(t) = x(t) + iH(x(t))$, where $H(x(t))$ is the Hilbert transform of $x(t)$, providing the signal's quadrature component. The instantaneous amplitude of the signal is then given by $A_x(t) = |z(t)| = \sqrt{x(t)^2 + H(x(t))^2}$ and the instantaneous phase is computed as $\phi_x(t) = \tan^{-1}\left(\frac{\mathrm{Im}(z(t))}{\mathrm{Re}(z(t))}\right)$. For two signals $x(t)$ and $y(t)$, their instantaneous phase difference at each time point is defined as: $\Delta\phi(t) = \phi_x(t) - \phi_y(t)$ PLI is then derived by calculating the expectation of the sign function applied to the sine of the phase difference [84]:

$$\mathrm{PLI} = |\mathbb{E}\left[\mathrm{sgn}(sin(\Delta\phi(t))\right]| \tag{6.1}$$

where $\mathbb{E}[\cdot]$ denotes expectation (or average) over time, sgn(.) represents sign function that returns $+1$ if the phase difference is positive, $-1$ if negative, and 0 if exactly zero, $sin(\Delta\phi(t))$ captures whether the phase difference at time $t$ is positive or negative. PLI quantifies the consistency of directional phase relationships by focusing on whether one signal consistently leads or lags the other. A PLI value

of 0 indicates no consistent phase lag, while a value of 1 reflects perfect phase synchronization with a consistent lag.

By excluding zero lag phase differences, PLI minimizes the influence of spurious connectivity caused by volume conduction or noise. Its normalized range between 0 and 1 makes it well suited for comparing connectivity patterns across different datasets and experimental conditions, establishing it as a robust measure for studying functional brain networks [38].

### Weighted Phase Lag Index

Weighted phase lag index enhances PLI by accounting for the magnitude of phase lag asymmetry, offering a more precise and sensitive measure of functional connectivity [85]. While PLI quantifies the consistency of nonzero phase lags to eliminate volume conduction and common source artifacts, it does not account for the strength of these phase lags. wPLI refines this approach by weighting the contribution of phase differences based on the magnitude of their imaginary components, enhancing its robustness, particularly in noisy or low signal conditions [85].

wPLI is derived from the instantaneous phases of the signals, which are extracted from their analytic representations, and is calculated as follows:

$$\text{wPLI} = \frac{|\mathbb{E}\left[\text{Im}(\Delta\phi(t)) \cdot \text{sgn}(\text{Im}(\Delta\phi(t)))\right]|}{\mathbb{E}\left[|\text{Im}(\Delta\phi(t))|\right]} \tag{6.2}$$

where $\text{Im}(\Delta\phi(t))$ is the imaginary part of the phase difference, and $\text{sgn}(\text{Im}(\Delta\phi(t)))$ is the sign function which determines whether the imaginary part is positive or negative [85]. The numerator quantifies the extent of phase lag asymmetry by incorporating its magnitude, while the denominator serves as a normalization factor, ensuring that wPLI values remain constrained within the range of 0 to 1. A wPLI value approaching 0 signifies the absence of a consistent phase lag, whereas values near 1 indicate strong phase synchronization with a stable directional lag.

By emphasizing nonzero phase lags, wPLI effectively reduces spurious connec-

tivity arising from volume conduction or noise, as these factors generally produce symmetric phase differences [86]. The bounded range of wPLI enables standardized comparisons across datasets and experimental conditions, establishing it as a robust and reliable metric for assessing functional brain networks.

### Debiased Weighted Phase Lag Index

Debiased weighted phase lag index is an advanced functional connectivity metric that builds upon the wPLI by further reducing statistical biases [85]. While wPLI improves upon PLI by weighting phase differences based on their magnitude, it remains susceptible to bias in the presence of a limited number of observations [14]. This bias can distort connectivity estimates, particularly when analyzing short time windows or low signal conditions. The dWPLI addresses this limitation by correcting for the estimation bias, thereby providing a more reliable and unbiased measure of functional connectivity, especially for small sample sizes. It estimates phase synchronization that addresses some of the limitations of traditional phase locking measures [83]. It is designed to reduce biases introduced by volume conduction effects, which can distort connectivity measures in EEG due to the spread of electrical activity across the scalp [87]. The dwPLI calculates the consistency of phase synchronization between two EEG signals, reflecting how much the phase difference between them is consistent over time [88]. By debiasing, dwPLI reduces the influence of volume conduction artifacts that can affect the accuracy of phase synchronization measures [87]. Equation 6.3 shows the mathematical representation of dwPLI.

$$\widehat{\Omega}^{\mathrm{w}}(\mathrm{f}) = \frac{\sum_{\mathrm{k}=1}^{\mathrm{K}} \sum_{\mathrm{j}\neq\mathrm{k}} \mathrm{im}\left[\mathrm{s}_{\mathrm{k}}(\mathrm{f})\right] \mathrm{im}\left[\mathrm{s}_{\mathrm{j}}(\mathrm{f})\right]}{\sum_{\mathrm{k}=1}^{\mathrm{K}} \sum_{\mathrm{j}\neq\mathrm{k}} \left|\mathrm{im}\left[\mathrm{s}_{\mathrm{k}}(\mathrm{f})\right] \mathrm{im}\left[\mathrm{s}_{\mathrm{j}}(\mathrm{f})\right]\right|} \tag{6.3}$$

### Coherence

Coherence is a fundamental measure in functional connectivity analysis, quantifying the degree of correlation between the phase of oscillatory activity in two EEG

signals across different frequency bands. It provides insight into the synchrony of neural oscillations between spatially distinct brain regions, thereby serving as an important tool for investigating neural communication and network dynamics [89]. Coherence is mathematically defined as the squared magnitude of the cross spectral density between two signals, normalized by the product of their individual power spectral densities:

$$C_{rs}(f) = \frac{|P_{rs}(f)|^2}{P_{rr}(f)\,P_{ss}(f)} \tag{6.4}$$

where $C_{rs}(f)$ represents the coherence between two signals $r(t)$ and $s(t)$ at a given frequency $f$. $P_{rs}(f)$ denotes the cross spectral density, and $P_{rr}(f)$ and $P_{ss}(f)$ correspond to the power spectral densities of the individual signals. Coherence values range between 0 and 1, where a coherence value close to 1 signifies strong phase synchronization, indicating robust functional connectivity between the two brain regions. A coherence value approaching 0 suggests weak or absent phase synchronization, potentially reflecting disrupted communication or independent neural activity in the respective regions.

### Amplitude Envelope Correlation

Amplitude envelope correlation quantifies the statistical dependencies between the amplitude envelopes of oscillatory brain activity [90]. Unlike phase-based connectivity measures, which assess the synchrony of oscillation phases, AEC evaluates the correlation between the temporal fluctuations of the amplitude envelopes of neural signals. This approach provides insight into the strength of neural interactions while being less susceptible to confounding effects such as volume conduction and zero lag correlations [90].

The amplitude envelope $A_f(t)$, which represents the instantaneous power of the oscillation, can be computed as $A_f(t) = |z_f(t)| = \sqrt{x_f(t)^2 + H(x_f(t))^2}$, where $H(x_f(t))$ is the Hilbert transform of filtered input signal. Based on this, amplitude envelopes of two EEG signals, $A_x(t)$ and $A_y(t)$, extracted from different brain

regions, are then correlated using the Pearson correlation coefficient to quantify their coupling and is represented using equation 6.5.

$$\text{AEC} = \frac{\mathbb{E}[(A_x(t) - \mu_{A_x})(A_y(t) - \mu_{A_y})]}{\sigma_{A_x}\sigma_{A_y}} \tag{6.5}$$

where $\mathbb{E}[\cdot]$ denotes the expectation (time average), $\mu_{A_x}$ and $\mu_{A_y}$ are the mean amplitude envelopes, $\sigma_{A_x}$ and $\sigma_{A_y}$ are standard deviations.

AEC captures slow fluctuations in signal power, which are often linked to large scale brain network activity, such as in resting-state studies. Unlike phase-based measures like wPLI, AEC is unaffected by phase noise, focusing instead on the co-fluctuation of signal intensities. This makes AEC particularly valuable for studying frequency dependent connectivity across brain networks.

### 6.1.2 Graph Theory Metrics

Graph theory provides a robust mathematical framework for modeling and analyzing complex brain networks, facilitating the examination of functional and structural connectivity patterns [39]. By representing the brain as a graph where nodes correspond to brain regions and edges represent connections (functional or structural), graph theoretical metrics can quantify key aspects of brain organization. These metrics provide insights into information flow, network integration, and segregation, which are crucial for understanding brain function under normal and pathological conditions. In particular, they are instrumental in assessing disruptions in neural communication observed in disorders of consciousness and other neurocognitive impairments [91].

#### Clustering Coefficient

The clustering coefficient quantifies the extent to which nodes in a network form local clusters, representing the degree of inter-connectivity among neighboring

nodes [44]. Mathematically, the clustering coefficient $C_i$ of a node $i$ is defined as:

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \tag{6.6}$$

**Path Length**

The path length refers to the number of edges in the shortest path between two nodes, representing the efficiency of information transfer in a network. The shortest path length between two nodes $i$ and $j$ is denoted as $d_{ij}$. The characteristic path length of the entire network is computed as the average shortest path length between all pairs of nodes:

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \tag{6.7}$$

where $N$ is the total number of nodes. A shorter path length indicates that information can travel rapidly across the network, which is crucial for cognitive functions such as perception, decision making, and motor responses. In brain networks, longer path lengths may indicate disrupted connectivity, which is often observed in coma, vegetative states, and neurodegenerative conditions [14].

**Small-World Topology**

Small-world networks exhibit a balance between local specialization and global integration, characterized by high clustering coefficients and short path lengths. This topology enables efficient information processing with minimal wiring cost, which is a fundamental property of biological and neural systems [91]. The small-worldness index $S$ is calculated as:

$$S = \frac{C/C_{\text{rand}}}{L/L_{\text{rand}}} \tag{6.8}$$

where $C$ and $L$ are the clustering coefficient and characteristic path length of the observed brain network, $C_{rand}$ and $L_{rand}$ are the corresponding values in a randomly generated network with the same number of nodes and edges. A network is considered small-world if $S > 1$, meaning that it has higher clustering than a random network while maintaining a similar or slightly higher path length. Small-world properties support efficient information transfer across the brain by allowing for both segregated processing in specialized modules and rapid integration of information across distant regions. Disruptions in small-world topology have been observed in neurological disorders, including Alzheimer's disease, schizophrenia, and disorders of consciousness, where the balance between local and global information processing is impaired [91].

#### Modularity

Modularity is a fundamental graph theoretical measure that quantifies the degree to which a network can be divided into distinct communities or modules. In the context of brain networks, modularity reflects the tendency of functionally related brain regions to form tightly interconnected clusters, while maintaining fewer connections with other modules. This organization supports efficient local processing within modules and optimized global communication across the entire network [44].

### 6.1.3 Sliding Window Correlation

Sliding window correlation (SWC) analysis is a technique used in analyzing resting-state EEG signals to study dynamic functional connectivity patterns in brain [92]. This method involves computing time dependent correlations between EEG channels by segmenting the continuous EEG signal into overlapping or non-overlapping time windows and calculating correlation metrics, such as Pearson's correlation coefficient, within each window. The sliding window approach captures temporal variations in

functional connectivity, revealing transient network states that static correlation methods may overlook. By continuously shifting the window across the EEG signal, researchers can construct time varying connectivity matrices that illustrate how brain network interactions evolve over time. This method is particularly useful in understanding neural dynamics associated with cognitive states, neurological disorders, and brain plasticity [93]. However, selecting an appropriate window length is crucial, as shorter windows enhance temporal resolution but may introduce spurious correlations, while longer windows improve stability but may smooth out transient network changes.

### 6.1.4 Canonical Correlation Analysis

Canonical correlation analysis is a statistical method designed to analyze and quantify the relationships between two sets of multivariate data [94]. This technique is valuable when exploring associations within complex datasets, such as brain network metrics and clinical outcomes. In this research, CCA is utilized to evaluate the prognostic relevance of three brain networks, each characterized by six graph metrics, and their correlation with GOSE scores recorded three months after the initial assessment. By applying CCA, the research aims to uncover how different brain network attributes, derived from multiple connectivity measures, relate to long-term clinical outcomes [95].

## 6.2 Choice of ANN Models for Multiclass Classification

The proposed work employs several advanced artificial neural network architectures to classify states of consciousness coma, UWS, MCS, and healthy subjects using dynamic functional connectivity features, namely wPLI and AEC. The architectures include three variations of MLP, simple recurrent neural network, long short-term

memory network, and gated recurrent units, each tailored to optimize the learning process for different data representations and dynamics.

### 6.2.1 MLP Models

MLP model 1 is initialized with an input layer designed to accept flattened input data with a shape of (X_train.shape[1]), facilitating computational efficiency. The architecture includes a fully connected hidden layer consisting of 128 neurons with ReLU activation, followed by dropout regularization at a rate of 0.5 to mitigate overfitting. A subsequent hidden layer incorporates 64 neurons, also employing ReLU activation, with dropout regularization at a rate of 0.3 for additional regularization. The output layer consists of four neurons (corresponding to four classes) and utilizes a softmax activation function to generate class probabilities. The model is optimized using the Adam algorithm with default hyperparameters, and categorical cross-entropy is employed as the loss function, appropriate for multi-class classification tasks. The training process spans 50 epochs with a batch size of 32, utilizing 10% of the training dataset for validation.

MLP model 2 extends the architecture of MLP model 1 by incorporating an additional fully connected hidden layer with 256 neurons and ReLU activation, followed by batch normalization to enhance training stability and convergence speed. A dropout rate of 0.4 is applied to this layer. The second hidden layer comprises 128 neurons with ReLU activation, batch normalization, and dropout at a rate of 0.4. The third hidden layer contains 64 neurons, also employing ReLU activation and dropout at a rate of 0.3. To further combat overfitting, L2 kernel regularization (penalty = 0.001) is applied to the second and third hidden layers. The output layer remains consistent with MLP model 1, utilizing a softmax activation function with four units. Optimization is conducted using the Adam algorithm with a learning rate of 0.001, and the training parameters are identical to those of MLP model 1.

MLP model 3 follows a similar structure to MLP model 2 but adopts a stream-

lined architecture to optimize learning efficiency. It comprises two hidden layers with 256 and 128 neurons, respectively, both employing ReLU activation, batch normalization, dropout at a rate of 0.4, and L2 regularization (penalty = 0.001). The output layer is identical to those of the previous models, and the Adam optimizer is utilized. Early stopping and learning rate reduction callbacks are implemented to terminate training when validation performance ceases to improve, thereby optimizing computational resource utilization.

### 6.2.2 Sequential Models: Simple RNN, LSTM, and GRU

The simple RNN model is designed to process sequential data, necessitating the reshaping of input data into the format (time_steps, features). The first recurrent layer consists of 128 simple RNN units with tanh activation, ensuring the retention of outputs for subsequent layers (return_sequences=True). Dropout regularization is applied at a rate of 0.4. A second recurrent layer, comprising 64 simple RNN units, is configured with return_sequences=False and employs dropout at a rate of 0.3. Batch normalization is applied after each layer to stabilize gradient propagation. Following the RNN layers, a dense layer with 32 units and rectified linear unit (ReLU) activation is incorporated, along with dropout regularization at a rate of 0.2. The output layer consists of four softmax activated neurons. The model is optimized using the Adam algorithm with a learning rate of 0.0005, employing categorical cross-entropy as the loss function. Performance evaluation metrics include accuracy, precision, recall, and F1-score.

The LSTM model leverages the ability of long short-term memory layers to capture long-term dependencies in sequential data. Similar to the simple RNN model, the input data is reshaped to (time_steps, features). The first LSTM layer consists of 128 units with return_sequences=True, applying dropout regularization at a rate of 0.4 to mitigate overfitting. The second LSTM layer comprises 64 units with return_sequences=False and dropout regularization at a rate of 0.3.

Batch normalization is applied following each layer to enhance training stability. The post-LSTM dense layers include a fully connected layer with 32 units, ReLU activation, and dropout at a rate of 0.2, followed by an output layer with four softmax activated neurons. Optimization is conducted using the Adam algorithm with a learning rate of 0.0005, incorporating early stopping and learning rate reduction callbacks to ensure efficient training.

The GRU model follows an architectural structure similar to the LSTM model but replaces LSTM layers with GRUs, which are computationally more efficient while preserving the capability to model sequential dependencies. The first GRU layer comprises 128 units with return_sequences=True and dropout regularization at a rate of 0.4. The second GRU layer consists of 64 units with return_sequences=False and dropout regularization at a rate of 0.3. Batch normalization is applied following each GRU layer. The dense layers include a 32-unit fully connected layer with ReLU activation and dropout at a rate of 0.2, culminating in an output layer with four softmax-activated neurons. Like the LSTM model, the GRU model employs the Adam optimizer (learning rate = 0.0005), categorical cross-entropy loss, and early stopping to optimize training performance.

The hybrid CNN-LSTM model integrates convolutional and recurrent architectures to enhance sequential learning capabilities. Input data is structured as (time_steps, num_features) to capture temporal dependencies effectively. The model begins with two Conv1D layers comprising 64 and 128 filters, utilizing ReLU activation. These layers are followed by MaxPooling1D layers to downsample the feature maps. Batch normalization is applied after each convolutional layer to stabilize the training process. The extracted spatial features are processed through two LSTM layers with 64 and 32 units, employing tanh activation to model sequential dependencies. Dropout regularization is applied at rates of 0.3 and 0.2 within the LSTM layers to mitigate overfitting. A fully connected dense layer, consisting of 32 neurons with ReLU activation, is followed by dropout at a rate of 0.2. The final

output layer utilizes a softmax activation function for multi-class classification. The model is trained using categorical cross-entropy as the loss function and optimized using the Adam algorithm with a learning rate of 0.0005. Training is conducted with a batch size of 32 over 50 epochs, incorporating early stopping and a learning rate reduction scheduler to enhance convergence and overall performance.

## 6.3 Methodology

### 6.3.1 Functional Connectivity based Classification using Deep Learning Models

The deep learning based classification of DOC cohorts were performed using PLI, wPLI, and AEC features extracted from the preprocessed signal. The preprocessed EEG signal consists of 19 channels: F3, F4, C3, C4, F7, F8, P3, P4, T7, T8, P7, P8, O1, O2, Fp1, Fp2, Fz, Pz, and Cz. Each channel is analyzed across five frequency bands: delta, theta, alpha, beta, and gamma. Consequently, each connectivity feature results in a symmetric 19×19 feature matrix for each frequency band. Due to the symmetry of these matrices, only the upper triangular elements, excluding the diagonal, are utilized for classification. This results in 171 unique elements per frequency band. With five frequency bands, the total feature vector for each time window contains 855 values (5 × 171). For a 20 second window size, each subject generates 95 windows, yielding a feature matrix of size (95, 855) per subject. With 15 subjects per group, the total number of samples per group is 1425 (15 × 95). Considering four groups, the dataset comprises 5700 samples (4 × 1425), each represented by a feature vector of dimensionality 855. After reshaping the data into a 2D array, the final input matrix has dimensions of (5700, 855). For a 19 second window, each subject generates 103 windows, resulting in a feature matrix of size (103, 855) per subject. With 15 subjects per group, the total number of samples per group is 1545 (15 × 103). For four groups, the dataset contains 6180

samples (4 × 1545). For an 18 second window, each subject produces 109 windows, leading to a final dataset of 6540 feature vectors (109 × 15 × 4). For a 17 second window, each subject generates 115 windows, resulting in a dataset of 6900 feature vectors. For a 16 second window, each subject produces 120 windows, yielding 7200 feature vectors in total. Figure 6.1 describes the proposed methodology adopted for the FC-based multiclass classification of DOC patients using various deep learning models.

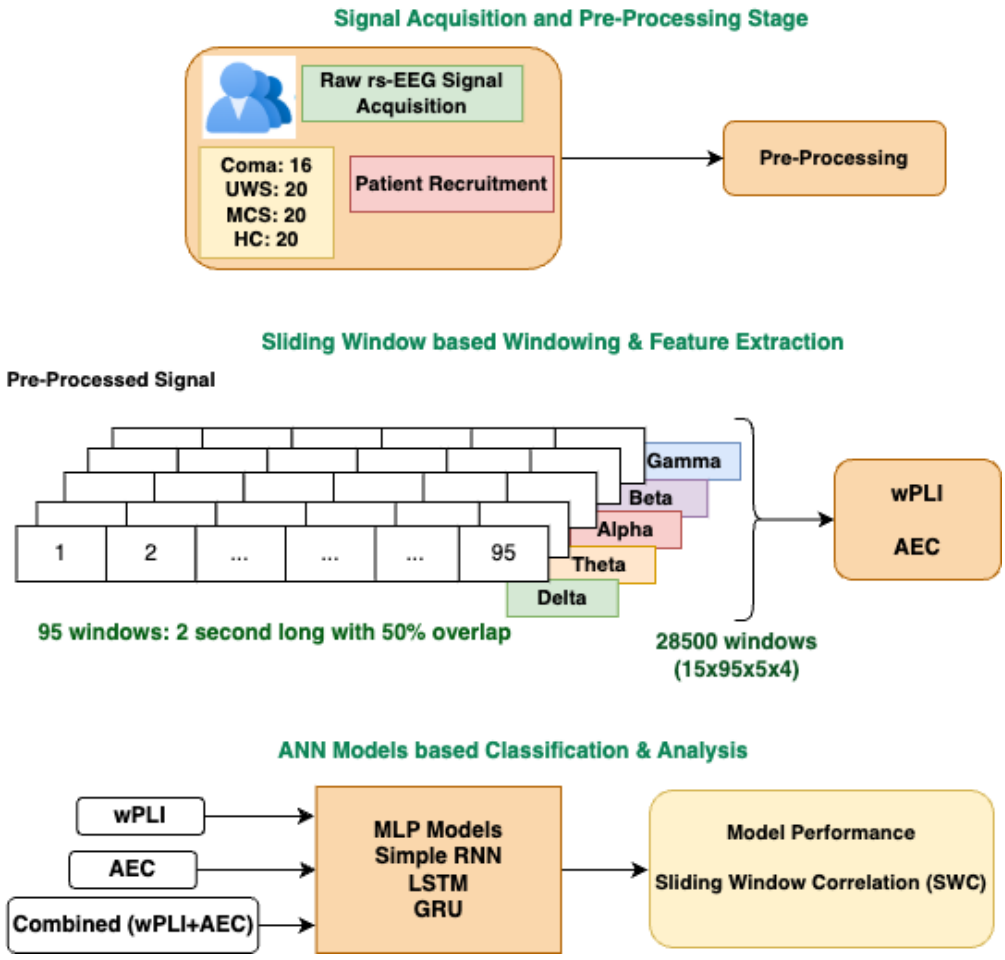



**Figure 6.1:** Proposed methodology: FC-based multiclass classification of DOC patients using various deep learning models

### 6.3.2 Graph Theoretical Analysis of Undirected Weighted Brain Networks

This section presents an analysis of functional connectivity patterns through the construction of undirected, weighted brain networks derived from three distinct connectivity measures namely AEC, dwPLI, and Coherence. Network-based features are extracted to characterize connectivity differences across various DOC states. The predictive potential of these features is assessed using CCA in relation to GOSE scores obtained three months following the initial assessment.

## 6.4 Results

### 6.4.1 Performance Evaluation of Deep Learning Models

To assess the effectiveness of deep learning approaches in classifying different states of consciousness, a series of artificial neural network models were trained using functional connectivity features derived from EEG data. These models aimed to differentiate between coma, UWS, MCS, and healthy controls based on connectivity measures such as PLI, dwPLI, and AEC.

Table 6.1 presents a comparative analysis of six neural network models comprising three MLP architectures, a simple RNN, an LSTM, and a GRU based on performance metrics including accuracy, precision, recall, and F1-score. These models were trained to classify consciousness states using dynamic functional connectivity features derived from PLI, dwPLI, and AEC.

The performance metrics highlighted how variations in both model architecture and feature type influence the classification of subjects into coma, UWS, MCS, and healthy control groups. The results demonstrated the nuanced capacity of different models to leverage temporal brain connectivity features for effective state discrimination. Distinct trends across models and feature sets underscored the importance of selecting appropriate neural network structures and functional

**Table 6.1:** Performance comparison of different ANN models for multiclass classification with 20 second window length and 80/20 train-test split (coma vs. UWS vs. MCS vs. HC) using three different features and a combination of two of them.

| **Metrics** | **Feature** | **MLP Model 1** | **MLP Model 2** | **MLP Model 3** | **Simple RNN** | **LSTM** | **GRU** | **Hybrid CNN-LSTM** |
|---|---|---|---|---|---|---|---|---|
| Accuracy | PLI | 0.548 | 0.622 | 0.618 | 0.52 | 0.533 | 0.529 | 0.571 |
| Precision | | 0.553 | 0.626 | 0.619 | 0.529 | 0.543 | 0.559 | 0.572 |
| Recall | | 0.548 | 0.622 | 0.618 | 0.52 | 0.533 | 0.529 | 0.571 |
| F1-Score | | 0.550 | 0.624 | 0.619 | 0.524 | 0.537 | 0.544 | 0.571 |
| Accuracy | wPLI | 0.772 | 0.532 | 0.515 | 0.668 | 0.638 | 0.713 | 0.48 |
| Precision | | 0.773 | 0.613 | 0.706 | 0.671 | 0.646 | 0.718 | 0.55 |
| Recall | | 0.772 | 0.532 | 0.515 | 0.668 | 0.638 | 0.713 | 0.48 |
| F1-Score | | 0.773 | 0.570 | 0.596 | 0.669 | 0.642 | 0.716 | 0.513 |
| Accuracy | AEC | 0.879 | 0.957 | 0.941 | 0.905 | 0.906 | 0.903 | 0.927 |
| Precision | | 0.879 | 0.958 | 0.942 | 0.907 | 0.908 | 0.904 | 0.928 |
| Recall | | 0.879 | 0.957 | 0.941 | 0.905 | 0.906 | 0.903 | 0.927 |
| F1-Score | | 0.879 | 0.958 | 0.942 | 0.906 | 0.907 | 0.903 | 0.928 |
| Accuracy | wPLI & AEC | 0.874 | 0.951 | 0.948 | 0.902 | 0.918 | 0.921 | 0.925 |
| Precision | | 0.876 | 0.952 | 0.948 | 0.904 | 0.918 | 0.921 | 0.925 |
| Recall | | 0.874 | 0.951 | 0.948 | 0.902 | 0.918 | 0.921 | 0.925 |
| F1-score | | 0.875 | 0.951 | 0.948 | 0.903 | 0.918 | 0.921 | 0.925 |

connectivity measures to optimize classification accuracy and clinical relevance in the assessment of disorders of consciousness.

#### 6.4.1.1 Performance Analysis on PLI Features

For PLI feature, MLP model 2 demonstrated the highest overall performance across all evaluated metrics, achieving an accuracy of 0.622, precision of 0.626, recall of 0.622, and an F1-score of 0.624. These results indicated that MLP model 2 effectively balances precision and recall, leading to the highest F1-score among all models assessed. In comparison, MLP models 1 and 3 exhibited slightly lower performance, with MLP model 3 outperforming MLP model 1 by a small margin.

The RNN-based models showd lower classification performance. The simple

RNN model achieved an accuracy of 0.52, precision of 0.529, recall of 0.52, and an F1-score of 0.524. Among recurrent architectures, the GRU model demonstrated marginally superior performance compared to the simple RNN and LSTM models, as reflected in its F1-score of 0.544, suggesting a slightly improved balance between precision and recall. The LSTM model also outperformed the simple RNN model slightly, yielding accuracy, precision, recall, and F1-score values of 0.533, 0.543, 0.533, and 0.537, respectively.

Table 6.2 summarizes the accuracy of different models utilizing PLI as a functional connectivity feature across various window sizes (ranging from 16 to 20 seconds) under an 80-20 fixed train-test split. The findings revealed that increasing the window size (particularly to 19 and 20 seconds) leads to improved model performance, especially for MLP-based architectures. MLP model 2 consistently attained the highest accuracy, peaking at 0.622 for 20 second window, while MLP model 3 follows closely with an accuracy of 0.6183. The recurrent models simple RNN, LSTM, and GRU exhibited relatively lower performance. However, both LSTM and GRU models showed an increasing trend in accuracy with longer window sizes, reaching their highest performance at 20 seconds (0.5325 and 0.5292, respectively).

**Table 6.2:** Accuracy with PLI features for different window sizes with 80-20 fixed train-test split

| **Model** | **16s** | **17s** | **18s** | **19s** | **20s** |
|---|---|---|---|---|---|
| **MLP Model 1** | 0.5376 | 0.5513 | 0.5547 | 0.5594 | 0.5479 |
| **MLP Model 2** | 0.6156 | 0.5974 | 0.6128 | 0.6197 | 0.622 |
| **MLP Model 3** | 0.6111 | 0.5958 | 0.607 | 0.6063 | 0.6183 |
| **Simple RNN** | 0.5165 | 0.5058 | 0.5083 | 0.5076 | 0.52 |
| **LSTM** | 0.4735 | 0.4722 | 0.4797 | 0.4688 | 0.5325 |
| **GRU** | 0.4874 | 0.4684 | 0.4622 | 0.4782 | 0.5292 |
| **Hybrid CNN-LSTM** | 0.5502 | 0.5585 | 0.5608 | 0.561 | 0.5705 |

Table 6.3 presents the classification results obtained using group k-fold cross-validation (CV) with PLI-based connectivity features. The results further validated

the superior performance of MLP-based models over RNN-based architectures, with MLP model 2 achieving the highest accuracy and F1-score across all window sizes. An increase in window size corresponds with improved classification outcomes, particularly for MLP models, suggesting their effectiveness in capturing long-range temporal dependencies.

The evaluation across different cross-validation folds highlighted that folds 8 and 10 yield the most favorable results, particularly for MLP models 2 and 3. This suggested that these data splits offer a more optimal distribution for training and testing. Recurrent models (simple RNN, LSTM, and GRU) exhibited greater performance variability and consistently underperformed relative to MLP architectures. The overall results suggested that MLP models, particularly models 2 and 3, were better suited for leveraging PLI-based features, likely due to their capacity for learning high-dimensional and complex representations more effectively than sequential models.

**Table 6.3:** Best fold obtained using group k-fold cross validation and the performance metrics of each model using PLI features

| **16s** | | | **17s** | | **18s** | | **19s** | | **20s** | |
|---|---|---|---|---|---|---|---|---|---|---|
| **MLP M1** CV: 10 | Acc. | 0.5545 | CV: 5 | 0.5633 | CV: 8 | 0.5646 | CV: 8 | 0.5727 | CV: 8 | 0.5779 |
| | Pr. | 0.5608 | | 0.5653 | | 0.567 | | 0.5775 | | 0.5822 |
| | ReC. | 0.5545 | | 0.5633 | | 0.5646 | | 0.5727 | | 0.5779 |
| | F1-S | 0.5576 | | 0.5642 | | 0.5658 | | 0.5751 | | 0.58 |
| **MLP M2** CV: 10 | Acc. | 0.6151 | CV: 8 | 0.6261 | CV: 10 | 0.6258 | CV: 8 | 0.6339 | CV: 10 | 0.64 |
| | Pr. | 0.616 | | 0.627 | | 0.6263 | | 0.6347 | | 0.6408 |
| | ReC. | 0.6151 | | 0.6261 | | 0.6258 | | 0.6339 | | 0.64 |
| | F1-S | 0.6474 | | 0.6596 | | 0.6582 | | 0.6665 | | 0.6730 |
| **MLP M3** CV: 10 | Acc. | 0.6145 | CV: 8 | 0.6264 | CV: 10 | 0.6283 | CV: 8 | 0.6339 | CV: 8 | 0.6418 |
| | Pr. | 0.6152 | | 0.6275 | | 0.6289 | | 0.6345 | | 0.6425 |
| | ReC. | 0.6145 | | 0.6264 | | 0.6283 | | 0.6339 | | 0.6418 |
| | F1-S | 0.6145 | | 0.6271 | | 0.6299 | | 0.6342 | | 0.6431 |
| **Simple RNN** CV: 10 | Acc. | 0.5218 | CV: 10 | 0.5238 | CV: 8 | 0.5275 | CV: 10 | 0.5376 | CV: 8 | 0.5365 |
| | Pr. | 0.5271 | | 0.5296 | | 0.5347 | | 0.5449 | | 0.5462 |
| | ReC. | 0.5218 | | 0.5238 | | 0.5275 | | 0.5376 | | 0.5365 |
| | F1-S | 0.4819 | | 0.4799 | | 0.4850 | | 0.4970 | | 0.4933 |
| **LSTM** CV: 5 | Acc. | 0.4807 | CV: 10 | 0.491 | CV: 5 | 0.4861 | CV: 8 | 0.4914 | CV: 8 | 0.4866 |
| | Pr. | 0.4869 | | 0.4955 | | 0.4912 | | 0.4984 | | 0.4892 |
| | ReC. | 0.4807 | | 0.491 | | 0.4861 | | 0.4914 | | 0.4866 |
| | F1-S | 0.4641 | | 0.4774 | | 0.4688 | | 0.4728 | | 0.4642 |
| **GRU** CV: 10 | Acc. | 0.4996 | CV: 10 | 0.4953 | CV: 5 | 0.5 | CV: 8 | 0.5076 | CV: 5 | 0.502 |
| | Pr. | 0.5065 | | 0.4997 | | 0.5058 | | 0.5111 | | 0.5059 |
| | ReC. | 0.4996 | | 0.4953 | | 0.5 | | 0.5076 | | 0.502 |
| | F1-S | 0.5127 | | 0.4996 | | 0.5099 | | 0.5176 | | 0.5118 |
| **Hybrid CNN-LSTM** CV: 10 | Acc. | 0.562 | CV:8 | 0.5709 | CV:8 | 0.5711 | CV:10 | 0.5763 | CV:8 | 0.5822 |
| | Pr. | 0.5629 | | 0.5735 | | 0.572 | | 0.577 | | 0.5832 |
| | ReC. | 0.562 | | 0.5709 | | 0.5711 | | 0.5763 | | 0.5822 |
| | F1-S | 0.5624 | | 0.5722 | | 0.5715 | | 0.5766 | | 0.5827 |

#### 6.4.1.2 Performance Analysis on wPLI Features

For the wPLI feature, which characterizes phase-based functional connectivity between brain regions, model performance varied considerably across architectures. Among all evaluated models, MLP model 1 demonstrated the highest classification performance, achieving an accuracy of 0.772. This strong performance is further supported by a precision of 0.773, indicating its capability to correctly classify positive cases while minimizing false positives. The recall score of 0.772 suggested that the model effectively captures a substantial proportion of true positives, while its F1-score of 0.773 highlights a well-balanced trade-off between precision and recall. These findings underscored the ability of MLP model 1 to extract meaningful patterns from the wPLI feature set, thereby achieving high classification accuracy and generalizability.

In contrast, MLP models 2 and 3 exhibited significantly lower performance. MLP model 2 attains an accuracy of only 0.532, with corresponding precision, recall, and F1-score values of 0.613, 0.532, and 0.570, respectively, highlighting its difficulty in effectively generalizing across the dataset. MLP model 3 performed even worse, achieving an accuracy of 0.515. The poor performance of these models suggested that they struggle to effectively capture the complex phase-based connectivity information encoded within the wPLI feature set.

The simple RNN model exhibited moderate classification performance, with an accuracy of 0.668. While it does not achieve the same level of accuracy as MLP model 1 or the GRU model, it significantly outperformed MLP models 2 and 3. The model's precision (0.671) and recall (0.668) remain well-balanced, and its F1-score of 0.669 further confirmed its relatively stable classification capability.

The LSTM model, a more advanced recurrent neural network, achieved an accuracy of 0.638. Although this is lower than that of the simple RNN, its precision (0.646), recall (0.638), and F1-score (0.642) indicated a stable performance across metrics. However, the LSTM model underperformed compared to the GRU model,

which consistently outperformed it across all metrics. The comparatively lower performance of LSTM could be attributed to its increased architectural complexity, which may require a larger dataset and extensive hyperparameter tuning to achieve optimal performance. This observation highlighted the trade-off between model complexity and learning efficiency, particularly in scenarios with limited data.

Among the recurrent models, the GRU network emerged as the most effective. With an accuracy of 0.713, precision of 0.718, recall of 0.713, and F1-score of 0.716, GRU model demonstrated strong generalization capabilities and a well-balanced classification performance. The high F1-score suggested that the model effectively balances precision and recall, making it a reliable choice for classification tasks where both metrics are of equal importance. The superior performance of GRU compared to LSTM is likely due to its simpler architecture, which enabled efficient pattern recognition while mitigating the risk of overfitting and reducing computational demands. Better precision achieved by the GRU model suggested a reduced incidence of false positives, enhancing its suitability for applications where misclassification has significant consequences.

A comparative evaluation of all models revealed that MLP model 1 achieved the highest classification accuracy, effectively capturing meaningful connectivity patterns from the wPLI feature set. Meanwhile, the GRU model demonstrated the most consistent and well-balanced performance across all evaluation metrics. This makes GRU the most effective model for leveraging the temporal characteristics of the wPLI dataset while maintaining robust generalization capabilities.

Table 6.4 presents the accuracy of various deep learning models using wPLI features across multiple window sizes (ranging from 16 to 20 seconds) under an 80-20 fixed train-test split. The results indicate that increasing the window size enhanced accuracy for both the GRU and LSTM models. However, MLP model 1 remaind the most robust architecture, consistently achieving high classification performance across all window sizes.

**Table 6.4:** Accuracy with wPLI features for different window sizes with 80-20 fixed train-test split

| **Model** | **16s** | **17s** | **18s** | **19s** | **20s** |
|---|---|---|---|---|---|
| **MLP Model 1** | 0.7535 | 0.7649 | 0.7694 | 0.7625 | 0.7719 |
| **MLP Model 2** | 0.5086 | 0.5119 | 0.5772 | 0.5604 | 0.5319 |
| **MLP Model 3** | 0.4811 | 0.5265 | 0.5751 | 0.4979 | 0.5151 |
| **Simple RNN** | 0.6428 | 0.6388 | 0.6977 | 0.6819 | 0.6679 |
| **LSTM** | 0.5729 | 0.6175 | 0.5538 | 0.6157 | 0.6377 |
| **GRU** | 0.61 | 0.6296 | 0.6072 | 0.6382 | 0.7132 |
| **Hybrid CNN-LSTM** | 0.4 | 0.445 | 0.465 | 0.475 | 0.48 |

Table 6.5 presents the classification performance of models trained using wPLI-based features under a group k-fold cross-validation framework. The results highlight that MLP model 1 consistently outperformed all other models, attaining the highest accuracy of 0.785 at a 20 second window using fold 10. This finding suggested that MLP model 1 is particularly effective at capturing relevant connectivity patterns across different training splits. The simple RNN model also exhibited strong performance, reaching an accuracy of 0.7171 at a 20-second window, with fold 8 yielding the most favorable results in several cases. Among the MLP architectures, model 2 achieved its highest accuracy of 0.5798 at 20 seconds (fold 8), while model 3 reaches a peak accuracy of 0.5703 at 20 seconds (fold 5). However, these models displayed lower recall and F1-scores compared to MLP model 1, suggesting inferior generalization capabilities.

**Table 6.5:** Best fold obtained using group k-fold cross validation and the performance metrics of each model using wPLI features

| | **16s** | | **17s** | | **18s** | | **19s** | | **20s** | |
|---|---|---|---|---|---|---|---|---|---|---|
| **MLP M1** CV: 10 | Acc. | 0.7622 | CV: 10 | 0.7713 | CV: 8 | 0.7769 | CV: 10 | 0.7827 | CV: 10 | 0.785 |
| | Pr. | 0.7631 | | 0.7728 | | 0.7781 | | 0.7839 | | 0.786 |
| | ReC. | 0.7622 | | 0.7713 | | 0.7769 | | 0.7827 | | 0.785 |
| | F1-S | 0.7626 | | 0.7720 | | 0.7775 | | 0.7833 | | 0.7855 |
| **MLP M2** CV: 8 | Acc. | 0.5414 | CV: 5 | 0.5509 | CV: 10 | 0.5566 | CV: 10 | 0.561 | CV: 8 | 0.5798 |
| | Pr. | 0.6639 | | 0.643 | | 0.6597 | | 0.6652 | | 0.6759 |
| | ReC. | 0.5414 | | 0.5509 | | 0.5566 | | 0.561 | | 0.5798 |
| | F1-S | 0.5964 | | 0.5934 | | 0.6038 | | 0.6087 | | 0.6242 |
| **MLP M3** CV: 10 | Acc. | 0.5399 | CV: 10 | 0.5661 | CV: 10 | 0.5558 | CV: 8 | 0.5672 | CV: 5 | 0.5703 |
| | Pr. | 0.6499 | | 0.6651 | | 0.6578 | | 0.6643 | | 0.6379 |
| | ReC. | 0.5399 | | 0.5661 | | 0.5558 | | 0.5672 | | 0.5703 |
| | F1-S | 0.5898 | | 0.6116 | | 0.6025 | | 0.6119 | | 0.6022 |
| **Simple RNN** CV: 8 | Acc. | 0.6773 | CV: 10 | 0.7015 | CV: 8 | 0.7048 | CV: 10 | 0.7106 | CV: 8 | 0.7171 |
| | Pr. | 0.6804 | | 0.7042 | | 0.7066 | | 0.7123 | | 0.7189 |
| | ReC. | 0.6773 | | 0.7015 | | 0.7048 | | 0.7106 | | 0.7171 |
| | F1-S | 0.6788 | | 0.7028 | | 0.7057 | | 0.7114 | | 0.7180 |
| **LSTM** CV: 8 | Acc. | 0.6131 | CV: 8 | 0.6293 | CV: 10 | 0.6256 | CV: 8 | 0.6378 | CV:10 | 0.6317 |
| | Pr. | 0.616 | | 0.6326 | | 0.6276 | | 0.6416 | | 0.6359 |
| | ReC. | 0.6131 | | 0.6293 | | 0.6256 | | 0.6378 | | 0.6317 |
| | F1-S | 0.6145 | | 0.6309 | | 0.6266 | | 0.6397 | | 0.6338 |
| **GRU** CV: 5 | Acc. | 0.614 | CV: 10 | 0.6405 | CV: 8 | 0.6383 | CV: 8 | 0.6515 | CV: 10 | 0.639 |
| | Pr. | 0.6198 | | 0.6449 | | 0.6421 | | 0.6549 | | 0.6443 |
| | ReC. | 0.614 | | 0.6405 | | 0.6383 | | 0.6515 | | 0.639 |
| | F1-S | 0.6169 | | 0.6427 | | 0.6402 | | 0.6532 | | 0.6416 |
| **Hybrid CNN-LSTM** CV: 10 | Acc. | 0.4706 | CV: 5 | 0.4621 | CV: 10 | 0.4828 | CV: 5 | 0.4874 | CV: 8 | 0.4935 |
| | Pr. | 0.527 | | 0.5173 | | 0.5463 | | 0.5642 | | 0.5723 |
| | ReC. | 0.4706 | | 0.4621 | | 0.4828 | | 0.4874 | | 0.4935 |
| | F1-S | 0.4972 | | 0.4881 | | 0.5126 | | 0.5230 | | 0.530 |

Recurrent models, such as LSTM and GRU, showed varying performance across different cross-validation folds. The LSTM model achieved its highest accuracy of 0.6378 at 19 seconds (fold 8), whereas the GRU model attained its peak accuracy of 0.6515 at 19 seconds (fold 8). These results indicated that the performance of recurrent models is highly dependent on the specific training splits used during cross-validation.

#### 6.4.1.3 Performance Analysis on AEC Features

Utilizing the AEC feature set, MLP model 2 demonstrated the highest overall classification performance, achieving an accuracy of 95.7%, precision of 95.8%, recall of 95.7%, and an F1-score of 95.8%. These results indicated the model's exceptional capacity to correctly classify instances, minimize false positives, and effectively identify true positives. The F1-score, which represents the harmonic mean of precision and recall, further confirmed that MLP model 2 maintains a well-balanced performance without favoring one metric over the other. The model's ability to efficiently extract meaningful patterns from the AEC feature set is likely attributable to its optimized architecture, which enabled it to capture intricate interdependencies among features.

MLP model 3 also exhibited strong classification performance, attaining an accuracy of 94.1%, precision of 94.2%, recall of 94.1%, and an F1-score of 94.2%. Although its performance is slightly lower than that of MLP model 2, it effectively generalized across the dataset and demonstrated a robust ability to process the AEC feature set. The marginal performance gap between MLP models 2 and 3 may be attributed to architectural differences that make model 3 slightly less optimal for capturing fine-grained patterns in the data.

MLP model 1 achieved an accuracy of 87.9%, precision of 87.9%, recall of 87.9%, and an F1-score of 87.9%. While its performance does not reach the levels achieved by MLP models 2 and 3, it still demonstrated reliable classification capabilities.

The model successfully identified patterns within the AEC feature set, although its comparatively simpler architecture may impose limitations on its capacity to capture complex relationships within the data.

The simple RNN model achieved a strong classification performance, with an accuracy of 90.5%, precision of 90.7%, recall of 90.5%, and an F1-score of 90.6%. These findings indicate that, despite its recurrent architecture primarily being designed for sequential data, the simple RNN model generalized effectively when applied to the AEC feature set. While its performance surpassed that of MLP model 1, it remained inferior to MLP models 2 and 3, suggesting that the simple RNN does not fully exploit the potential of AEC-based connectivity features as efficiently as the MLP architectures.

The LSTM model demonstrated performance comparable to that of the simple RNN, achieving an accuracy of 90.6%, precision of 90.8%, recall of 90.6%, and an F1-score of 90.7%. These results suggested that, while LSTM is a powerful model for sequential data processing, its increased architectural complexity does not confer a substantial advantage when applied to static connectivity features such as AEC.

The GRU model achieved an accuracy of 90.3%, precision of 90.4%, recall of 90.3%, and an F1-score of 90.3%. While its performance is closely aligned with that of the simple RNN and LSTM, its slightly lower values suggested that it may be marginally less suited for processing AEC-based features compared to the other recurrent models.

Table 6.6 presents the accuracy of various deep learning models utilizing AEC features across different window sizes under an 80-20 fixed train-test split. The results indicated that MLP model 2 consistently achieved high accuracy across all evaluated window sizes, maintaining values above 95%, thereby emerging as the most effective model for AEC-based classification. Among recurrent models, the simple RNN outperformed both the LSTM and GRU architectures, achieving its

best accuracy at a 17 second window size.

**Table 6.6:** Accuracy with AEC features for different window sizes with 80-20 fixed train-test split

| **Model** | **16s** | **17s** | **18s** | **19s** | **20s** |
|---|---|---|---|---|---|
| **MLP Model 1** | 0.8851 | 0.8864 | 0.8765 | 0.8838 | 0.8788 |
| **MLP Model 2** | 0.9626 | 0.9546 | 0.9624 | 0.9561 | 0.9572 |
| **MLP Model 3** | 0.9514 | 0.9486 | 0.952 | 0.943 | 0.9412 |
| **Simple RNN** | 0.9049 | 0.9126 | 0.909 | 0.9026 | 0.9047 |
| **LSTM** | 0.8535 | 0.8467 | 0.8388 | 0.8303 | 0.9063 |
| **GRU** | 0.8506 | 0.8599 | 0.8524 | 0.8503 | 0.9028 |
| **Hybrid CNN-LSTM** | 0.925 | 0.9293 | 0.9272 | 0.9242 | 0.9273 |

The group k-fold cross-validation results for models trained using AEC-based features, presented in Table 6.7, revealed that MLP models consistently achieved higher accuracy than other architectures, with MLP models 2 and 3 demonstrating the best performance. MLP model 2 attained a peak accuracy of 0.9648 at a 20 second window size (fold 8), followed closely by MLP model 3, which achieved an accuracy of 0.9648 at 19 seconds (fold 10). These results highlight the superior ability of MLP models to leverage AEC-based features, suggesting that they effectively capture the underlying patterns within the data. While MLP model 1 also performed well, reaching a maximum accuracy of 0.8895 at 16 seconds (fold 8), it is consistently outperformed by MLP models 2 and 3.

**Table 6.7:** Best fold obtained using group k-fold cross validation and the performance metrics for each model using AEC features.

| | | 16s | | 17s | | 18s | | 19s | | 20s | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **MLP M1** CV: 8 | Acc. | 0.8895 | CV: 8 | 0.8885 | CV: 8 | 0.8839 | CV: 8 | 0.8876 | CV: 8 | 0.8834 |
| | Pr. | 0.8933 | | 0.8916 | | 0.8883 | | 0.891 | | 0.8869 |
| | ReC. | 0.8895 | | 0.8885 | | 0.8839 | | 0.8876 | | 0.8834 |
| | F1-S | 0.8914 | | 0.8900 | | 0.8861 | | 0.8893 | | 0.8851 |
| **MLP M2** CV: 10 | Acc. | 0.964 | CV: 8 | 0.9644 | CV: 8 | 0.9636 | CV: 8 | 0.9642 | CV: 8 | 0.9648 |
| | Pr. | 0.9645 | | 0.9651 | | 0.9644 | | 0.9648 | | 0.9654 |
| | ReC. | 0.964 | | 0.9644 | | 0.9636 | | 0.9642 | | 0.9648 |
| | F1-S | 0.9642 | | 0.9647 | | 0.9640 | | 0.9645 | | 0.9651 |
| **MLP M3** CV: 10 | Acc. | 0.9639 | CV: 10 | 0.9639 | CV: 8 | 0.964 | CV: 8 | 0.9648 | CV: 10 | 0.9622 |
| | Pr. | 0.9644 | | 0.9645 | | 0.9644 | | 0.9653 | | 0.9631 |
| | ReC. | 0.9639 | | 0.9639 | | 0.964 | | 0.9648 | | 0.9622 |
| | F1-S | 0.9641 | | 0.9642 | | 0.9642 | | 0.9650 | | 0.9626 |
| **Simple RNN** CV: 10 | Acc. | 0.916 | CV: 5 | 0.9192 | CV: 8 | 0.9225 | CV: 5 | 0.9189 | CV: 10 | 0.9212 |
| | Pr. | 0.9169 | | 0.9197 | | 0.9233 | | 0.9198 | | 0.9223 |
| | ReC. | 0.916 | | 0.9192 | | 0.9225 | | 0.9189 | | 0.9212 |
| | F1-S | 0.9164 | | 0.9194 | | 0.9229 | | 0.9193 | | 0.9217 |
| **LSTM** CV: 8 | Acc. | 0.8576 | CV: 8 | 0.8595 | CV: 8 | 0.8567 | CV: 10 | 0.8564 | CV: 10 | 0.8587 |
| | Pr. | 0.8584 | | 0.8602 | | 0.8575 | | 0.8573 | | 0.8594 |
| | ReC. | 0.8576 | | 0.8595 | | 0.8567 | | 0.8564 | | 0.8587 |
| | F1-S | 0.8580 | | 0.8598 | | 0.8571 | | 0.8568 | | 0.8590 |
| **GRU** CV: 8 | Acc. | 0.8641 | CV: 10 | 0.8645 | CV: 5 | 0.8686 | CV: 5 | 0.8627 | CV: 8 | 0.8705 |
| | Pr. | 0.8649 | | 0.8651 | | 0.8694 | | 0.863 | | 0.8716 |
| | ReC. | 0.8641 | | 0.8645 | | 0.8686 | | 0.8627 | | 0.8705 |
| | F1-S | 0.8645 | | 0.8648 | | 0.8690 | | 0.8628 | | 0.8710 |
| **Hybrid CNN-LSTM** CV: 10 | Acc. | 0.9283 | CV: 8 | 0.933 | CV: 5 | 0.9324 | CV: 5 | 0.9293 | CV: 10 | 0.9315 |
| | Pr. | 0.9292 | | 0.9336 | | 0.933 | | 0.9304 | | 0.9321 |
| | ReC. | 0.9283 | | 0.933 | | 0.9324 | | 0.9293 | | 0.9315 |
| | F1-S | 0.9287 | | 0.9332 | | 0.9326 | | 0.9297 | | 0.9317 |

Among RNN-based architectures, the simple RNN demonstrates the best performance, achieving an accuracy of 0.9225 at 18 seconds (fold 5), indicating its ability to effectively utilize AEC features. The GRU and LSTM models follow, with GRU attaining a peak accuracy of 0.8705 at 20 seconds (fold 8) and LSTM reaching a maximum accuracy of 0.8595 at 17 seconds (fold 8). These findings suggest that, while RNN-based models perform reasonably well with AEC-based features, they do not surpass the accuracy achieved by MLP models.

#### 6.4.1.4 Performance Analysis on Combined Features

MLP model 2 demonstrated the highest performance among all evaluated models, achieving accuracy, precision, recall, and F1-score values of 95.1%, 95.2%, 95.1%, and 95.1%, respectively. The high accuracy value indicated the model's strong capability in distinguishing between classes, while the closely aligned precision and recall values suggest that it effectively minimized both false positives and false negatives. However, these values were marginally lower than those obtained when utilizing AEC alone as the feature set.

MLP model 3 performed slightly below MLP model 2, attaining identical values of 94.8% for accuracy, precision, recall, and F1-score. These results indicated that the model is well-suited for capturing the complex relationships between wPLI and AEC features, albeit with slightly reduced performance compared to MLP model 2.

MLP model 1 exhibited a noticeable decline in performance, with an accuracy of 87.4%, precision of 87.6%, recall of 87.4%, and an F1-score of 87.5%. While it demonstrated the ability to process the combined wPLI and AEC feature set, the relatively lower performance suggested that its simpler architecture could limit its ability to fully capture the intricate relationships between these two types of connectivity features.

Among the recurrent models, the GRU architecture achieved high performance, attaining identical values of 92.1% for accuracy, precision, recall, and F1-score.

This suggested that GRU effectively captures the patterns within the combined wPLI-AEC feature set. The LSTM model performed slightly worse, with accuracy, precision, recall, and F1-score values of 91.8%. The simple RNN model, while achieving an accuracy of 90.2%, precision of 90.4%, recall of 90.2%, and an F1-score of 90.3%, performed moderately well but falls short in comparison to both the other recurrent architectures and the MLP models.

Table 6.8 presents the accuracy of various models utilizing a combination of AEC and wPLI across different window sizes (16–20 seconds) under an 80-20 train-test split. The MLP-based models outperformed the recurrent models, with MLP model 2 consistently achieving superior accuracy across all window sizes. The simple RNN model demonstrated better performance than LSTM and GRU in most cases, though GRU achieved high accuracy of 0.9209 at 20 seconds, surpassing LSTM. The best performance was observed at 18 seconds for MLP models and the simple RNN, while LSTM and GRU achievde peak performance at 20 seconds.

**Table 6.8:** Accuracy with AEC+wPLI features for different window sizes with 80-20 train test split

| **Model** | **16s** | **17s** | **18s** | **19s** | **20s** |
|---|---|---|---|---|---|
| **MLP Model 1** | 0.9083 | 0.8972 | 0.9167 | 0.9042 | 0.8735 |
| **MLP Model 2** | 0.9685 | 0.9686 | 0.955 | 0.9639 | 0.9511 |
| **MLP Model 3** | 0.9636 | 0.9671 | 0.9587 | 0.9584 | 0.9475 |
| **Simple RNN** | 0.9143 | 0.9099 | 0.9257 | 0.9076 | 0.9019 |
| **LSTM** | 0.8697 | 0.87 | 0.8564 | 0.8735 | 0.9177 |
| **GRU** | 0.8708 | 0.8771 | 0.8402 | 0.8756 | 0.9209 |
| **Hybrid CNN-LSTM** | 0.9015 | 0.915 | 0.9175 | 0.9155 | 0.925 |

Table 6.9 presents the best fold results obtained using group k-fold cross-validation for different models trained on the combined AEC and wPLI feature set. The results vary based on the evaluation metrics. Among the models, MLP Model 1 demonstrated the highest overall performance, achieving an accuracy of 0.9453 in its best fold (CV: 10). It also recorded higher precision (0.9463), recall (0.9453), and F1-score (0.9458). MLP model 2 followed closely, attaining an accuracy of 0.9383 (CV: 8), while MLP model 3 achieved a slightly lower peak accuracy of 0.928 (CV: 8). Among the recurrent models, GRU outperformed both the simple RNN and LSTM, achieving a peak accuracy of 0.8581 (CV: 10) and an F1-score of 0.8597, making it the most effective RNN-based model. The simple RNN and LSTM models exhibited relatively lower accuracy, with the simple RNN achieving a peak accuracy of 0.7518 (CV: 5) and LSTM reaching 0.756 (CV: 8). These findings further emphasized that MLP models significantly outperform RNN-based architectures in this experimental setup. Smaller window sizes (16–18 seconds) tend to yield better performance, particularly for MLP models, which consistently achieved higher accuracy within this range. GRU models exhibited stable performance across different window sizes but achieved their highest accuracy at 16 seconds, suggesting that this window size provides the most informative features for the GRU architecture. Recurrent models such as simple RNN and LSTM displayed better variability in performance, with LSTM achieving its best accuracy at 16 seconds and the simple RNN demonstrating slightly better results at 17 seconds and 20 seconds.

**Table 6.9:** Best performing fold in group k-fold cross validation and the performance metrics for each model using combined AEC and wPLI features.

| | | **16s** | | **17s** | | **18s** | | **19s** | | **20s** | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **MLP M1** CV: 10 | Acc. | 0.9453 | CV: 10 | 0.9107 | CV: 10 | 0.9437 | CV: 10 | 0.9434 | CV: 8 | 0.9118 |
| | Pr. | 0.9463 | | 0.9136 | | 0.9456 | | 0.9449 | | 0.9133 |
| | ReC. | 0.9453 | | 0.9107 | | 0.9437 | | 0.9434 | | 0.9118 |
| | F1-S | 0.9458 | | 0.9121 | | 0.9447 | | 0.9441 | | 0.9125 |
| **MLP M2** CV: 10 | Acc. | 0.9324 | CV: 10 | 0.9051 | CV: 10 | 0.9184 | CV: 10 | 0.9336 | CV: 8 | 0.9383 |
| | Pr. | 0.9345 | | 0.9087 | | 0.9212 | | 0.9349 | | 0.9391 |
| | ReC. | 0.9324 | | 0.9051 | | 0.9184 | | 0.9336 | | 0.9383 |
| | F1-S | 0.9334 | | 0.9069 | | 0.9198 | | 0.9342 | | 0.9387 |
| **MLP M3** CV: 10 | Acc. | 0.9256 | CV: 8 | 0.9076 | CV: 10 | 0.9172 | CV: 8 | 0.928 | CV: 5 | 0.9109 |
| | Pr. | 0.9273 | | 0.9105 | | 0.9199 | | 0.9292 | | 0.9144 |
| | ReC. | 0.9256 | | 0.9076 | | 0.9172 | | 0.928 | | 0.9109 |
| | F1-S | 0.9264 | | 0.9090 | | 0.9185 | | 0.9286 | | 0.9126 |
| **Simple RNN** $CV : 10$ | Acc. | 0.7331 | CV: 5 | 0.7518 | CV: 10 | 0.7147 | CV: 8 | 0.7364 | CV: 8 | 0.7496 |
| | Pr. | 0.7403 | | 0.7555 | | 0.72 | | 0.7396 | | 0.7556 |
| | ReC. | 0.7331 | | 0.7518 | | 0.7147 | | 0.7364 | | 0.7496 |
| | F1-S | 0.7367 | | 0.7536 | | 0.7173 | | 0.7380 | | 0.7526 |
| **LSTM** CV: 8 | Acc. | 0.756 | CV: 10 | 0.7339 | CV: 10 | 0.7122 | CV: 5 | 0.7137 | CV: 8 | 0.7447 |
| | Pr. | 0.7616 | | 0.7374 | | 0.7152 | | 0.7212 | | 0.7497 |
| | ReC. | 0.756 | | 0.7339 | | 0.7122 | | 0.7137 | | 0.7447 |
| | F1-S | 0.7588 | | 0.7356 | | 0.7137 | | 0.7174 | | 0.7472 |
| **GRU** $CV : 10$ | Acc. | 0.8581 | CV: 10 | 0.8499 | CV: 5 | 0.855 | CV: 8 | 0.8576 | CV: 10 | 0.8449 |
| | Pr. | 0.8613 | | 0.8537 | | 0.8555 | | 0.858 | | 0.8465 |
| | ReC. | 0.8581 | | 0.8499 | | 0.855 | | 0.8576 | | 0.8449 |
| | F1-S | 0.8597 | | 0.8518 | | 0.8552 | | 0.8578 | | 0.8457 |
| **Hybrid CNN-LSTM** CV: 5 | Acc. | 0.9047 | CV: 10 | 0.9243 | CV: 10 | 0.9189 | CV: 5 | 0.9166 | CV: 10 | 0.9266 |
| | Pr. | 0.908 | | 0.9251 | | 0.9214 | | 0.9189 | | 0.9277 |
| | ReC. | 0.9047 | | 0.9243 | | 0.9189 | | 0.9166 | | 0.9266 |
| | F1-S | 0.9063 | | 0.9247 | | 0.9201 | | 0.9177 | | 0.9271 |

The 95% confidence intervals (CI) for precision, accuracy, F1-score, and recall were computed to assess the performance of various models across four feature sets: PLI, wPLI, AEC, and the combined AEC and wPLI. Among the evaluated models, MLP models 2 and 3 exhibited higher accuracy, particularly when utilizing AEC (0.964, 95% CI = 0.964–0.965). While MLP model 1 demonstrated slightly lower accuracy, its performance improved significantly when employing the combined AEC and wPLI feature set (0.931, 95% CI = 0.909–0.953).

The recurrent models (Simple RNN, LSTM, and GRU) displayed varying levels of performance, with GRU surpassing both LSTM and Simple RNN, especially when using the combined AEC and wPLI features (0.853, 95% CI = 0.846–0.86). The CNN-LSTM hybrid model achieved a competitive performance, attaining an accuracy of 0.931 (95% CI = 0.928–0.933) with AEC and 0.918 (95% CI = 0.908–0.929) with the combined AEC and wPLI feature set, making it comparable to the high performing MLP models.

Across all models, AEC consistently yielded high classification performance, while the combination of AEC with wPLI further enhanced accuracy, particularly for the MLP and GRU architectures. A comprehensive summary of the confidence intervals for the implemented models is provided in Table 6.10 and 6.11.

**Table 6.10:** Confidence interval for the implemented models using different features

| Model | Performance Metrics | PLI | wPLI | AEC | AEC+wPLI |
|---|---|---|---|---|---|
| **MLP Model 1** | Accuracy | 0.567 (0.555, 0.578) | 0.776 (0.764, 0.787) | 0.887 (0.883, 0.89) | 0.931 (0.909, 0.953) |
| | Precision | 0.571 (0.559, 0.582) | 0.777 (0.765, 0.788) | 0.89 (0.887, 0.893) | 0.933 (0.911, 0.955) |
| | Recall | 0.567 (0.555, 0.578) | 0.776 (0.764, 0.787) | 0.887 (0.883, 0.89) | 0.931 (0.909, 0.953) |
| | F1Score | 0.569 (0.557, 0.58) | 0.776 (0.765, 0.788) | 0.888 (0.885, 0.892) | 0.932 (0.91, 0.954) |
| **MLP Model 2** | Accuracy | 0.628 (0.617, 0.64) | 0.558 (0.54, 0.576) | 0.964 (0.964, 0.965) | 0.926 (0.909, 0.942) |
| | Precision | 0.629 (0.617, 0.641) | 0.662 (0.647, 0.676) | 0.965 (0.964, 0.965) | 0.928 (0.912, 0.943) |
| | Recall | 0.628 (0.617, 0.64) | 0.558 (0.54, 0.576) | 0.964 (0.964, 0.965) | 0.926 (0.909, 0.942) |
| | F1Score | 0.661 (0.649, 0.673) | 0.605 (0.59, 0.62) | 0.964 (0.964, 0.965) | 0.927 (0.91, 0.943) |
| **MLP Model 3** | Accuracy | 0.629 (0.616, 0.641) | 0.56 (0.544, 0.575) | 0.964 (0.963, 0.965) | 0.918 (0.907, 0.929) |
| | Precision | 0.63 (0.617, 0.642) | 0.655 (0.641, 0.669) | 0.964 (0.963, 0.965) | 0.92 (0.91, 0.93) |
| | Recall | 0.629 (0.616, 0.641) | 0.56 (0.544, 0.575) | 0.964 (0.963, 0.965) | 0.918 (0.907, 0.929) |
| | F1Score | 0.63 (0.617, 0.643) | 0.604 (0.592, 0.615) | 0.964 (0.963, 0.965) | 0.919 (0.909, 0.93) |

**Table 6.11:** Confidence interval for the implemented models using different features

| Model | Performance Metrics | PLI | wPLI | AEC | AEC+wPLI |
|---|---|---|---|---|---|
| **Simple RNN** | Accuracy | 0.529 (0.52, 0.538) | 0.702 (0.683, 0.721) | 0.92 (0.916, 0.923) | 0.737 (0.719, 0.756) |
| | Precision | 0.536 (0.526, 0.547) | 0.704 (0.686, 0.723) | 0.92 (0.917, 0.924) | 0.742 (0.724, 0.76) |
| | Recall | 0.529 (0.52, 0.538) | 0.702 (0.683, 0.721) | 0.92 (0.916, 0.923) | 0.737 (0.719, 0.756) |
| | F1Score | 0.487 (0.478, 0.497) | 0.703 (0.685, 0.722) | 0.92 (0.917, 0.923) | 0.74 (0.721, 0.758) |
| **LSTM** | Accuracy | 0.487 (0.482, 0.493) | 0.628 (0.616, 0.639) | 0.858 (0.856, 0.859) | 0.732 (0.708, 0.756) |
| | Precision | 0.492 (0.486, 0.498) | 0.631 (0.619, 0.643) | 0.859 (0.857, 0.86) | 0.737 (0.713, 0.761) |
| | Recall | 0.487 (0.482, 0.493) | 0.628 (0.616, 0.639) | 0.858 (0.856, 0.859) | 0.732 (0.708, 0.756) |
| | F1Score | 0.469 (0.462, 0.477) | 0.629 (0.617, 0.641) | 0.858 (0.857, 0.86) | 0.735 (0.711, 0.758) |
| **GRU** | Accuracy | 0.501 (0.495, 0.506) | 0.637 (0.62, 0.654) | 0.866 (0.862, 0.87) | 0.853 (0.846, 0.86) |
| | Precision | 0.506 (0.501, 0.511) | 0.641 (0.625, 0.657) | 0.867 (0.862, 0.871) | 0.855 (0.848, 0.862) |
| | Recall | 0.501 (0.495, 0.506) | 0.637 (0.62, 0.654) | 0.866 (0.862, 0.87) | 0.853 (0.846, 0.86) |
| | F1Score | 0.51 (0.502, 0.519) | 0.639 (0.622, 0.655) | 0.866 (0.862, 0.871) | 0.854 (0.847, 0.861) |
| **Hybrid CNN-LSTM** | Accuracy | 0.573, (0.563, 0.582) | 0.479, (0.463, 0.495) | 0.931, (0.928, 0.933) | 0.918, (0.908, 0.929) |
| | Precision | 0.574, (0.564, 0.583) | 0.545, (0.516, 0.575) | 0.932, (0.929, 0.934) | 0.92, (0.911, 0.93) |
| | Recall | 0.573, (0.563, 0.582) | 0.479, (0.463, 0.495) | 0.931, (0.928, 0.933) | 0.918, (0.908, 0.929) |
| | F1Score | 0.573, (0.564, 0.582) | 0.51, (0.489, 0.532) | 0.931, (0.929, 0.934) | 0.919, (0.909, 0.929) |

#### 6.4.1.5 Sliding Window Correlation Analysis

The sliding window correlation analysis of PLI, wPLI and AEC connectivity matrices provided valuable insights into the functional connectivity of brain and its relationship to varying levels of consciousness. By examining the temporal stability and variability of connectivity patterns across five frequency bands, this study highlighted the dynamic nature of neural interactions underlying different conscious states.

Two key metrics, mean SWC and standard deviation of SWC (Std. SWC), were employed to quantify these patterns. The mean SWC represents the overall consistency of connectivity across consecutive time windows, whereas Std. SWC captures the temporal variability of these connectivity patterns, offering a comprehensive perspective on network dynamics.

**SWC Analysis for PLI**

The analysis of mean SWC across different frequency bands revealed a progressive increase in functional connectivity from coma (0.0196 in delta) to healthy controls (0.0593 in delta), with similar trends observed in theta (0.0208 to 0.0540), alpha (0.0222 to 0.0602), beta (0.0363 to 0.0554), and gamma (0.0183 to 0.0586) bands. These findings indicated that higher levels of consciousness were associated with stronger and more dynamic brain network interactions. Healthy controls exhibited high mean SWC values across all frequency bands, reflecting robust connectivity dynamics, whereas coma patients demonstrated low values, suggesting diminished neural interactions. UWS and MCS groups occupied an intermediate position, with subtle differences depending on the frequency band, particularly in theta and beta. Std. SWC, which quantifies the variability of connectivity dynamics, followed a similar trend. Healthy individuals exhibited high variability across all frequency bands, with peak Std. SWC values in delta (0.1446), theta (0.1428), alpha (0.1431), beta (0.1381), and gamma (0.1447), indicating high complexity and

adaptability in their neural interactions. Coma patients displayed lower variability, suggesting more static connectivity dynamics, while UWS and MCS groups showed intermediate levels of variability, with overlapping patterns in theta and gamma bands. Notably, the increased heterogeneity observed in MCS could indicated the onset of more intricate neural interactions, distinguishing it from UWS.

**SWC Analysis for wPLI**

In delta band, the mean SWC increased markedly from coma (0.00003) to healthy controls (0.034), indicating a evident disparity in connectivity stability. Similar trends were observed in the theta band, with mean SWC values progressing from coma (0.0038) to UWS (0.0061), MCS (0.029), and healthy controls (0.054), suggesting a gradual restoration of stable network dynamics corresponding with higher levels of consciousness. This pattern extended to the alpha and beta bands, where coma patients showed negative or near-zero mean SWC values (-0.0025 in alpha and -0.0067 in beta), while healthy controls exhibited significantly higher values (0.035 in alpha and 0.034 in beta). In gamma band, although healthy controls presented a high mean SWC of 0.030, MCS group surpassed them with a value of 0.036, potentially indicating distinctive gamma band connectivity associated with partial recovery of consciousness. Coma patients displayed reduced gamma mean SWC (0.0075), signifying diminished high frequency network stability. Regarding variability in connectivity dynamics, as captured by the Std. SWC, healthy controls again showed high values across most frequency bands, indicating a flexible and adaptive neural architecture characteristic of healthy brain function. In theta and alpha bands, healthy controls exhibited Std. SWC values of 0.236 and 0.243, respectively, while coma patients recorded lower variability (0.178 in both bands). A similar progression was noted in beta band, where variability increased from coma (0.178) to healthy controls (0.241). In gamma band, while healthy controls exhibited the highest variability (0.244), MCS group followed closely (0.188), suggesting

partially restored dynamic flexibility in high frequency networks. The consistently lower Std. SWC in coma patients, particularly in gamma (0.175), underscored their impaired capacity for dynamic connectivity adaptation, aligning with the reduced complexity and rigidity of brain function observed in severely disordered states of consciousness.

**SWC Analysis for AEC**

High mean SWC values across all frequency bands were observed in healthy controls, indicating robust and stable functional connectivity dynamics in healthy brain. In delta band, healthy individuals demonstrated a mean SWC of 0.6487, markedly higher than those observed in coma (0.3093) and MCS (0.3098) patients, with UWS patients occupying an intermediate position (0.4406), suggesting a partial restoration of low frequency network stability. This trend persisted across other frequency bands. In theta band, associated with attention and memory, healthy controls maintained high mean SWC (0.6609), exceeding coma (0.3458), UWS (0.4119), and MCS (0.3271) values, reflecting the compromised capacity for stable cognitive processing in pathological groups. The alpha band, which supports sensory and resting-state processing, showed a similar gradation: healthy controls recorded a mean SWC of 0.6409, nearly double that of coma patients (0.3072), while UWS (0.3681) and MCS (0.3267) groups exhibited modest preservation. In beta band, which underpins higher order cognition, healthy controls dominated (0.6591), followed by UWS (0.3962), MCS (0.3264), and coma (0.3214). The gamma band, essential for neural synchronization and cognitive integration, showed healthy controls with a mean SWC of 0.6490, in contrast to substantially lower values in coma (0.3445) and MCS (0.3395) patients. Regarding variability in connectivity dynamics, measured by Std. SWC, healthy controls demonstrated moderate variability across all bands, consistently lower than their respective mean SWC values, indicating an optimal balance between stability and adaptability.

In delta band, healthy controls exhibited a Std. SWC of 0.1984, compared to higher variability in coma (0.2526) and MCS (0.2506), suggesting more erratic dynamics in impaired states. UWS patients showed intermediate variability (0.2396), aligning with their partial functional preservation. Similar patterns emerged in the theta and alpha bands, where healthy controls maintained lower variability (0.2064 and 0.2120, respectively) compared to pathological groups, whose elevated variability implies disorganized and unstable connectivity. In beta band, healthy individuals displayed the lowest variability (0.1952), while coma, UWS, and MCS show increasing variability (0.2378, 0.2664, and 0.2616, respectively), indicative of compromised network efficiency. The gamma band echoed this trend, with healthy controls at 0.2266 and higher values in coma (0.2394), UWS (0.2579), and MCS (0.2541), further emphasizing the diminished dynamic regulation of high frequency connectivity in disordered states of consciousness.

### 6.4.2 Graph Theoretical Network Analysis of Consciousness States

The proficiency of the extracted graph metrics in distinguishing between DOC groups was evaluated using the XGBoost classifier. The classifier achieved an accuracy of 73%, demonstrating the effectiveness of the metrics as features for categorizing patients based on their level of consciousness. The classification result underscores the potential of graph theory as a tool for understanding the functional organization of the brain in DOC patients.

#### 6.4.2.1 AEC-based Brain Network Analysis

Graph theoretical metrics derived from AEC-based networks exhibit distinct variations across the three DOC groups, as illustrated in Figure 6.2. Significant differences were evident in key network properties, including the clustering coefficient, modularity, and betweenness centrality. The clustering coefficient tends

to be higher in UWS patients than in coma and MCS, suggesting a shift toward more localized connectivity patterns. This trend implied that, despite the extensive disruption of global network communication, UWS patients may maintain more strongly clustered local connections.

Modularity, a measure of the brain's capacity to form distinct functional modules, varied notably across the DOC groups, highlighting fundamental differences in the brain's structural organization. These variations likely reflected alterations in functional segregation and integration associated with impaired consciousness. The observed patterns emphasized the interplay between localized processing and global network interactions, providing valuable insights into the neural mechanisms underlying different levels of consciousness impairment.

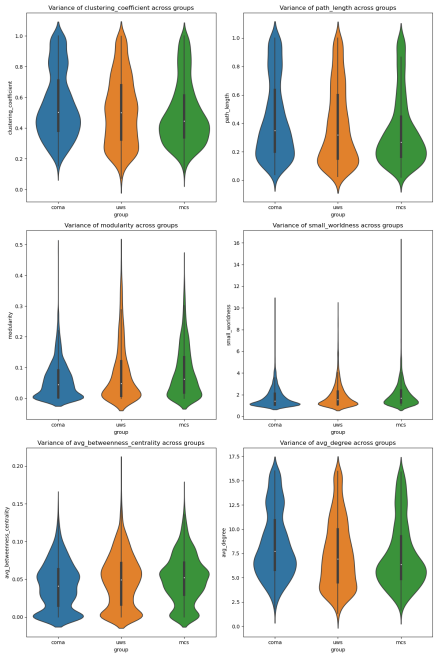



**Figure 6.2:** Variations of graph metrics from AEC network between groups

### 6.4.2.2 dwPLI-based Brain Network Analysis

Graph theoretical metrics derived from dwPLI-based networks demonstrate distinct trends across the DOC groups, as shown in Figure 6.3. Among the network properties analyzed, betweenness centrality and node degree emerged as particularly

informative features, offering insights into the centrality and influence of specific nodes within the brain network.

The analysis revealed that dwPLI-based networks effectively differentiate between UWS and MCS groups, as their network characteristics exhibit strong correlations with clinical outcome measures. These findings supported the premise that phase lag-based connectivity served as a reliable indicator of neural synchronization, which is critical for efficient communication across brain regions. The ability of dwPLI networks to capture subtle but meaningful distinctions reinforces their potential application in prognostic assessments of DOC patients.

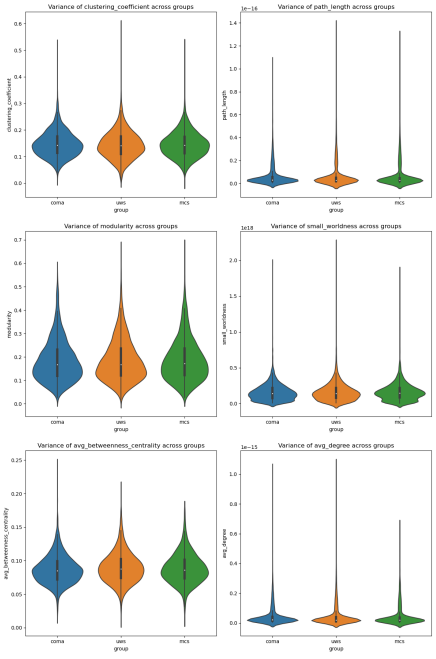



**Figure 6.3:** Variations of graph metrics from dwPLI network between groups

### 6.4.2.3 Coherence-Based Brain Network Analysis

Graph theoretical metrics derived from coherence-based networks further revealed significant differences among the DOC groups, as depicted in Figure 6.4. In particular, the clustering coefficient and small-worldness emerged as key distinguishing features. Small-worldness, which quantifies the balance between local clustering and global integration, is markedly higher in MCS patients compared to those in

coma and UWS. This suggested that MCS patients exhibit a more efficient network structure, potentially contributing to their relatively better cognitive function and awareness.

These findings aligned with the hypothesis that greater network integration and interconnectivity are associated with improved clinical outcomes in DOC patients. The increased small-worldness observed in MCS patients indicated that, despite impairments in overall consciousness, these individuals retain a more optimized network organization that facilitates neural communication. These observations enhanced the understanding of differential connectivity patterns across DOC groups and highlight the utility of graph theoretical analysis in evaluating and characterizing states of consciousness.

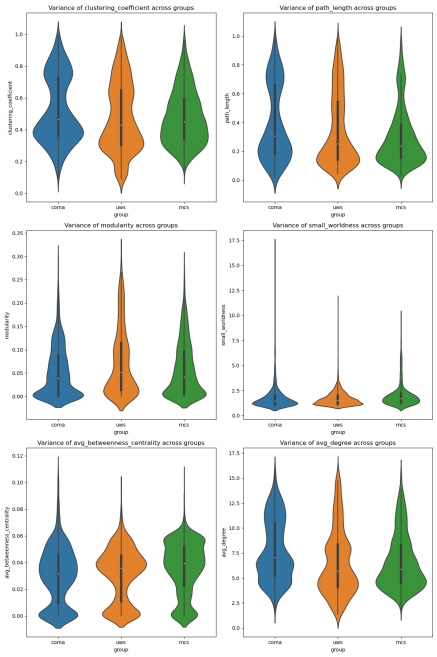



**Figure 6.4:** Variations of graph metrics from coherence network between groups

### Canonical Correlation Analysis

The canonical correlation analysis revealed distinct and nuanced relationships between brain network features and clinical outcomes, as measured by the GOSE scores. Among the examined networks, both the dwPLI-based and coherence-based

networks demonstrated high canonical correlations of 0.91, indicating a strong association between the corresponding brain metrics and recovery outcomes. This suggested that variations in these network features are strongly predictive of clinical recovery. However, despite similar correlation strengths, the directionality of influence differed between these networks. dwPLI network exhibited a positive canonical weight of 1, indicating that increases in dwPLI features were associated with more favorable outcomes. Coherence-based network displayed a negative canonical weight of -1, suggesting that elevated coherence values, which reflect increased synchrony in certain brain regions, may be indicative of poorer clinical recovery. These contrasting directions highlighted the complexity of functional connectivity in relation to consciousness and recovery, suggesting that while some connectivity patterns may facilitate recovery, others may reflect maladaptive network dynamics. Figure 6.5 illustrated the canonical weights associated with graph features derived from the dwPLI network.

The AEC-based network, which captures temporal coherence and synchrony of neural activity, yielded a slightly lower canonical correlation of 0.85. Although still strong, this value suggested a comparatively weaker association with GOSE scores relative to the dwPLI and coherence networks. The negative canonical weight observed for the AEC network implied that higher AEC feature values may correspond to lower GOSE scores, potentially indicating that increased activity or synchrony in specific neural regions is linked to less favorable recovery trajectories. The findings underscore the differential roles that distinct types of functional connectivity play in shaping clinical outcomes following severe brain injury.

## 6.5 Comparison with Existing Work

A comparison of the proposed results with existing techniques for differentiating states of disorders of consciousness is presented in Table 6.12. Among prior studies, Naro et al. [21] investigated the debiased weighted phase lag index as a phase-based

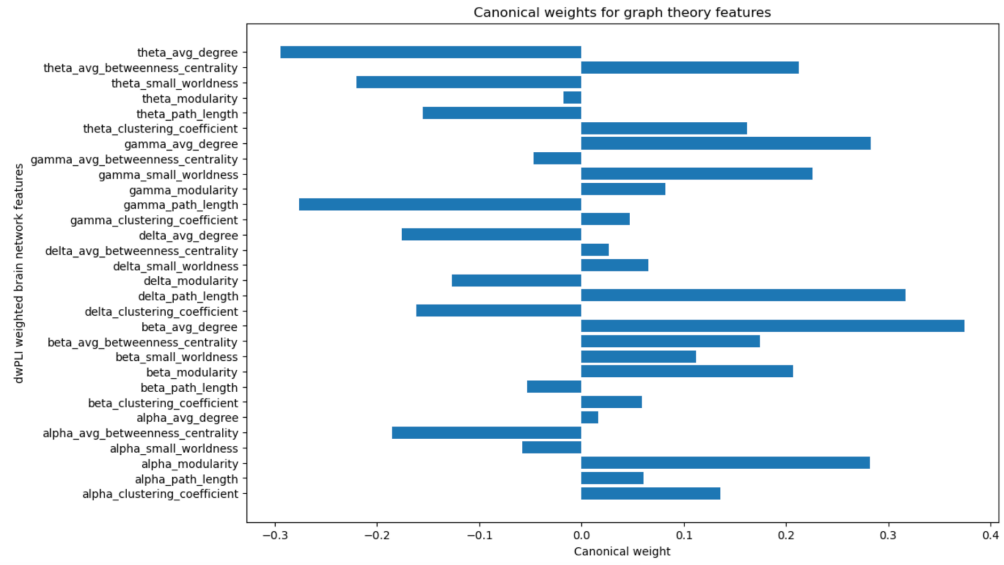



**Figure 6.5:** Canonical weights for graph theory features from dwPLI network

connectivity feature in a cohort of 15 MCS and 17 UWS patients. However, the absence of precision and accuracy hinders meaningful comparison and practical relevance of their findings.

Di Gregorio et al. [31] used phase coherence and dominant frequency features in combination with LDA classifier to differentiate DOC patients based on TBI status. Their study reported classification accuracies of 80% and 83.3% for the TBI and non-TBI groups, respectively, with corresponding precision scores of 0.777 and 0.857. While these results highlight the effectiveness of simple classifiers, linear techniques restricts how well they can capture the complex and nonlinear nature of brain dynamics.

In a related study, Duclos et al. (2021) [40] examined the use of AEC and wPLI to analyze anesthetic-induced unconsciousness and distinguish among different altered states of consciousness. Applying SVM classifier to data from nine control subjects, the study achieved classification accuracies of 69.4% using wPLI, 83.7% with AEC, and 84.1% when both features were combined. However, the small sample size limits the generalizability of these findings.

More recent approaches have leveraged ensemble learning techniques to improve adaptability. Yaqian Li et al. (2024) [96] utilized a CatBoost classifier, while

Sreelakshmi et al. (2024) [97] implemented a bagged tree model. Although these ensemble methods offer improved performance and robustness, they might fall short in flexibility compared to ANN models in handling heterogeneous and complex feature sets characteristic of EEG data.

**Table 6.12:** Comparison of our performance with the techniques in the literature. (TBI: traumatic brain injury)

| **Author** | **Features** | **No. of Subjects** | **Classifier** | **Accuracy (%)** | **Precision (%)** |
|---|---|---|---|---|---|
| Naro et al. (2018) [21] | dwPLI | MCS 15, UWS 17 | - | - | - |
| Di Gregorio et al. (2022) [31] | dominant frequency, PCoh, MI | DOC 33 | LDA | TBI : 80 ; Non-TBI: 83.3 | 77.7; 85.7 |
| Duglos et al. (2021) [40] | wPLI, AEC | controls 9 (anesthetic induced unconsciousness) | SVM | AEC: 83.7; wPLI: 69.4; AEC + wPLI: 84.1 | - |
| Yaqian Li et al. (2024) [96] | Microstate, dFC: MI | MCS: 16 UWS: 16 | CatBoost | 96.2 | - |
| Sreelakshmi et al. (2024) [97] | VMD mode-based features: kurtosis, skewness, spectral entropy, sample entropy | Coma 15, UWS 15, MCS 15 | Ensemble bagged tree | UWS vs. MCS: 83.3<br>Coma vs. UWS vs. MCS:76.0 | 84.2<br>78.9 |
| **Proposed Method** | **dFC feature: AEC** | **Coma 15, UWS 15, MCS 15, controls 15** | **ANN Models** | **95.7** | **95.8** |

## 6.6 Summary of the Chapter

This chapter investigated the diagnostic and prognostic potential of EEG-derived functional connectivity features in assessing levels of consciousness. Four connectivity measures namely PLI, dwPLI, AEC, and a combination of AEC and dwPLI were extracted from EEG signals and used to train several machine learning models. These included MLPs, RNNs, GRUs, LSTMs, and a CNN-LSTM hybrid architecture. The models were evaluated across multiple temporal window sizes using both 80-20 train-test split and group k-fold cross-validation. This chapter also employed a SWC analysis to explore the dynamic characteristics of functional connectivity of the 4 groups: UWS, coma, HC, and MCS. Healthy individuals consistently exhibited the highest mean SWC and the lowest Std. SWC across all frequency bands (delta, theta, alpha, beta, and gamma), suggesting stable and flexible network dynamics. Coma patients showed the lowest SWC values and least variability, indicating reduced and static connectivity. The UWS and MCS groups displayed intermediate characteristics, with MCS often exhibiting greater variability, potentially reflecting emerging neural complexity. CCA revealed strong associations between brain network features and clinical recovery outcomes, measured by the GOSE.

**The following research papers were published based on this chapter:**

1. Sreelakshmi Raveendran, et al., "Functional Connectivity in EEG: A Multiclass Classification Approach for Disorders of Consciousness" in *Frontiers in Neuroscience*, 2025, 19, 1550581.

2. Sreelakshmi Raveendran, et al., "Network Dynamics in Disorders of Consciousness: A Graph Theoretical Perspective" in *6th International Conference on Control, Communication and Computing (ICCC 2025)*, IEEE, Thiruvananthapuram, India, 2025.

# Chapter 7

# Conclusion and Future Work

This chapter summarizes the research work done in this thesis and provides directions for further research.

## 7.1 Conclusions

In the first part of this thesis, a comprehensive framework for classifying disorders of consciousness patients is presented, leveraging VMD as a preprocessing step to enhance EEG signal decomposition and feature extraction. This framework is built upon a dedicated dataset generated from resting-state EEG recordings of DOC patients and healthy controls at NIMHANS, ensuring standardized data acquisition and quality control. By applying VMD, the method aimed to reduce noise, minimize mode mixing, and extract physiologically relevant intrinsic mode functions, thereby improving the robustness of EEG-derived features. These refined features were then utilized for machine learning-based classification, enabling precise differentiation between DOC cohorts.

The second part of this thesis focused on functional connectivity and graph theory measures-based analysis of DOC cohorts to enhance both classification accuracy and interpretability. The incorporation of SHAP for feature selection has offered better insights into key EEG biomarkers associated with different

consciousness states, reinforcing the clinical relevance and transparency of AI-driven EEG analysis. A custom consciousness score equation was developed using the most significant EEG features identified through SHAP, aligning machine learning predictions with CRS-R scores. This custom equation enhanced the interpretability and clinical applicability of the proposed model by providing a quantitative measure of consciousness levels. These findings underscored the potential of artificial intelligence in DOC diagnosis, contributing to more accurate, reliable, and automated patient assessments while assisting clinicians in making informed decisions.

## 7.2 Future Directions

To further enhance model generalizability, future research will focus on expanding the dataset by incorporating a larger and more diverse population of DOC patients with additional resting-state EEG recordings. This will help address potential biases and improve model reliability across different patient groups. Multimodal analysis integrating EEG with other neuroimaging modalities, such as fMRI and magnetoencephalography, will be explored to provide a more comprehensive understanding of brain connectivity in DOC patients. Fusion of these modalities would uncover deeper neural patterns and enhance prognostic accuracy. Integration of vector quantization-based classification methods is proposed to improve feature representation and clustering, potentially leading to more refined and discriminative feature extraction. By adopting these advancements, future work aims to develop a more robust, interpretable, and clinically viable AI-driven framework for DOC assessment and prognosis prediction.

# List of Publications

## Journals

- Sreelakshmi Raveendran, Raghavendra Kenchiah, Santhos Kumar, Jayakrushna Sahoo, Farsana M K, Ravindranath Choudary, Sonia Bansal, Binu V S, A G Ramakrishnan, Subasree R, and Kala S. "Variational mode decomposition-based EEG analysis for the classification of disorders of consciousness," Frontiers in Neuroscience 18 (2024): 1340528. https://doi.org/10.3389/fnins.2024.1340528
- Sreelakshmi Raveendran, Kala S, Angarai Ganesan Ramakrishnan, Raghavendra Kenchiah, Jayakrushna Sahoo, Santhos Kumar, Farsana M K, Ravindranadh Chowdary Mundlamuri, Sonia Bansal, BINU V S, Subasree Ramakrishnan "Functional Connectivity in EEG: A Multiclass Classification Approach for Disorders of Consciousness", Frontiers in Neuroscience, 1550581. https://doi.org/10.3389/fnins.2025.1550581

## Conferences

- Sreelakshmi Raveendran, Raghavendra Kenchiah, Santhos Kumar, Jayakrushna Sahoo, Farsana M K, Ravindranath Choudary, Sonia Bansal, Binu V S, A G Ramakrishnan, Subasree R, and Kala S, "Scalp EEG-based Classification of Disorder of Consciousness States using Machine Learning Techniques," In 2023 11th International Symposium on Electronic Systems Devices and Computing (ESDC), vol. 1, pp. 1-6. IEEE, 2023. https://doi.org/10.1109/ESDC56251.2023.10149867
- Sreelakshmi Raveendran, Kala S, A G Ramakrishnan, Raghavendra Kenchiah, Jayakrushna Sahoo, Santhos Kumar, Farsana M K, Ravindranadh Chowdary Mundlamuri, Sonia Bansal, Binu V S, and Subasree R, "Network Dynamics

in Disorders of Consciousness: A Graph Theoretical Perspective", In 6th International Conference on Control, Communication and Computing (ICCC 2025), IEEE, Thiruvananthapuram, India, 2025

- Sreelakshmi Raveendran, Kala S, Raghavendra Kenchiah, Jayakrushna Sahoo, Farsana M K, Ravindranadh Chowdary Mundlamuri, Sonia Bansal, Binu V S, A G Ramakrishnan, and Subasree R, "Machine learning-based consciousness assessment in disorders of consciousness patients", In 2024 International Conference on Brain Computer Interface & Healthcare Technologies (iCon-BCIHT) (pp. 5-10). IEEE. https://doi.org/10.1109/iCon-BCIHT63907.2024.10882236

# References


[1] J. T. Giacino, J. J. Fins, S. Laureys, and N. D. Schiff, "Disorders of consciousness after acquired brain injury: the state of the science," *Nature Reviews Neurology*, vol. 10, no. 2, pp. 99–114, 2014.

[2] C. Schnakers, "Update on diagnosis in disorders of consciousness," *Expert review of neurotherapeutics*, vol. 20, no. 10, pp. 997–1004, 2020.

[3] Ç. G. Altıntop, F. Latifoğlu, and A. K. Akın, "Can patients in deep coma hear us? examination of coma depth using physiological signals," *Biomedical Signal Processing and Control*, vol. 77, p. 103756, 2022.

[4] A. Duszyk-Bogorodzka, M. Zieleniewska, and K. Jankowiak-Siuda, "Brain activity characteristics of patients with disorders of consciousness in the eeg resting state paradigm: a review," *Frontiers in systems neuroscience*, vol. 16, p. 654541, 2022.

[5] J. T. Giacino, "Disorders of consciousness: differential diagnosis and neuropathologic features," in *Seminars in neurology*, vol. 17, no. 02. © 1997 by Thieme Medical Publishers, Inc., 1997, pp. 105–111.

[6] N. Alnagger, P. Cardone, C. Martial, S. Laureys, J. Annen, and O. Gosseries, "The current and future contribution of neuroimaging to the understanding of disorders of consciousness," *La Presse Médicale*, vol. 52, no. 2, p. 104163, 2023.

[7] J. T. Giacino, D. I. Katz, N. D. Schiff, J. Whyte, E. J. Ashman, S. Ashwal, R. Barbano, F. M. Hammond, S. Laureys, G. S. Ling *et al.*, "Comprehensive systematic review update summary: disorders of consciousness: report of the guideline development, dissemination, and implementation subcommittee of the american academy of neurology; the american congress of rehabilitation medicine; and the national institute on disability, independent living, and rehabilitation research," *Neurology*, vol. 91, no. 10, pp. 461–470, 2018.

[8] C. Gervais, L.-P. Boucher, G. M. Villar, U. Lee, and C. Duclos, "A scoping review for building a criticality-based conceptual framework of altered states of consciousness," *Frontiers in Systems Neuroscience*, vol. 17, p. 1085902, 2023.

[9] B. Morris, J. Wong *et al.*, "A scoping review of treatments for the vegetative and minimally conscious states," *Brain Network and Modulation*, vol. 1, no. 2, p. 57, 2022.

[10] A. Thibaut, Y. G. Bodien, S. Laureys, and J. T. Giacino, "Minimally conscious state "plus": diagnostic criteria and relation to functional recovery," *Journal of neurology*, vol. 267, pp. 1245–1254, 2020.

[11] Y. Bai, Y. Lin, and U. Ziemann, "Managing disorders of consciousness: the role of electroencephalography," *Journal of neurology*, vol. 268, pp. 4033–4065, 2021.

[12] J. T. Giacino, C. Schnakers, D. Rodriguez-Moreno, K. Kalmar, N. Schiff, and J. Hirsch, "Behavioral assessment in patients with disorders of consciousness: gold standard or fool's gold?" *Progress in brain research*, vol. 177, pp. 33–48, 2009.

[13] J. T. Giacino, K. Kalmar, and J. Whyte, "The jfk coma recovery scale-revised: measurement characteristics and diagnostic utility," *Archives of physical medicine and rehabilitation*, vol. 85, no. 12, pp. 2020–2029, 2004.

[14] S. Chennu, P. Finoia, E. Kamau, J. Allanson, G. B. Williams, M. M. Monti, V. Noreika, A. Arnatkeviciute, A. Canales-Johnson, F. Olivares *et al.*, "Spectral signatures of reorganised brain networks in disorders of consciousness," *PLoS computational biology*, vol. 10, no. 10, p. e1003887, 2014.

[15] R. Jain and A. G. Ramakrishnan, "Electrophysiological and neuroimaging studies–during resting state and sensory stimulation in disorders of consciousness: a review," *Frontiers in neuroscience*, vol. 14, p. 555093, 2020.

[16] D. A. Engemann, F. Raimondo, J.-R. King, B. Rohaut, G. Louppe, F. Faugeras, J. Annen, H. Cassol, O. Gosseries, D. Fernandez-Slezak *et al.*, "Robust eeg-based cross-site and cross-protocol classification of states of consciousness," *Brain*, vol. 141, no. 11, pp. 3179–3192, 2018.

[17] S. Chennu, J. Annen, S. Wannez, A. Thibaut, C. Chatelle, H. Cassol, G. Martens, C. Schnakers, O. Gosseries, D. Menon *et al.*, "Brain networks predict metabolism, diagnosis and prognosis at the bedside in disorders of consciousness," *Brain*, vol. 140, no. 8, pp. 2120–2132, 2017.

[18] G. Teasdale and B. Jennett, "Assessment of coma and impaired consciousness: a practical scale," *The Lancet*, vol. 304, no. 7872, pp. 81–84, 1974.

[19] H. Gill-Thwaites, "The sensory modality assessment rehabilitation technique-a tool for assessment and treatment of patients with severe brain injury in a vegetative state," *Brain Injury*, vol. 11, no. 10, pp. 723–734, 1997.

[20] S. Corchs, G. Chioma, R. Dondi, F. Gasparini, S. Manzoni, U. Markowska-Kaczmar, G. Mauri, I. Zoppis, and A. Morreale, "Computational methods for resting-state EEG of patients with disorders of consciousness," *Frontiers in Neuroscience*, vol. 13, p. 807, 2019.

[21] A. Naro, A. Bramanti, A. Leo, A. Cacciola, A. Manuli, P. Bramanti, and R. S. Calabrò, "Shedding new light on disorders of consciousness diagnosis: the dynamic functional connectivity," *Cortex*, vol. 103, pp. 316–328, 2018.

[22] H. Lee, D. Golkowski, D. Jordan, S. Berger, R. Ilg, J. Lee, G. A. Mashour, U. Lee, M. S. Avidan, S. Blain-Moraes *et al.*, "Relationship of critical dynamics, functional connectivity, and states of consciousness in large-scale human brain networks," *Neuroimage*, vol. 188, pp. 228–238, 2019.

[23] S. Stefan, B. Schorr, A. Lopez-Rolon, I.-T. Kolassa, J. P. Shock, M. Rosenfelder, S. Heck, and A. Bender, "Consciousness indexing and outcome prediction with resting-state eeg in severe disorders of consciousness," *Brain Topography*, vol. 31, pp. 848–862, 2018.

[24] Y. Li, Y. Ran, M. Yao, and Q. Chen, "Altered static and dynamic functional connectivity of the default mode network across epilepsy subtypes in children: A resting-state fMRI study," *Neurobiology of Disease*, vol. 192, p. 106425, 2024.

[25] A. Naro, P. Bramanti, A. Leo, A. Cacciola, A. Bramanti, A. Manuli, and R. S. Calabrò, "Towards a method to differentiate chronic disorder of consciousness patients' awareness: The low-resolution brain electromagnetic tomography analysis," *Journal of the Neurological Sciences*, vol. 368, pp. 178–183, 2016.

[26] M. Amiri, P. M. Fisher, F. Raimondo, A. Sidaros, M. Cacic Hribljan, M. H. Othman, I. Zibrandtsen, S. S. Albrechtsen, O. Bergdal, A. E. Hansen *et al.*, "Multimodal prediction of residual consciousness in the intensive care unit: the connect-me study," *Brain*, vol. 146, no. 1, pp. 50–64, 2023.

[27] H. U. Amin, W. Mumtaz, A. R. Subhani, M. N. M. Saad, and A. S. Malik, "Classification of eeg signals based on pattern recognition approach," *Frontiers in computational neuroscience*, vol. 11, p. 103, 2017.

[28] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from eeg with deep recurrent-convolutional neural networks," *arXiv preprint arXiv:1511.06448*, 2015.

[29] R. Zafar, S. C. Dass, and A. S. Malik, "Electroencephalogram-based decoding cognitive states using convolutional neural network and likelihood ratio based score fusion," *PloS one*, vol. 12, no. 5, p. e0178410, 2017.

[30] S. Shanmugam and S. Dharmar, "A cnn-lstm hybrid network for automatic seizure detection in eeg signals," *Neural Computing and Applications*, vol. 35, no. 28, pp. 20 605–20 617, 2023.

[31] F. Di Gregorio, F. La Porta, V. Petrone, S. Battaglia, S. Orlandi, G. Ippolito, V. Romei, R. Piperno, and G. Lullini, "Accuracy of EEG biomarkers in the detection of clinical outcome in disorders of consciousness after severe acquired brain injury: preliminary results of a pilot study using a machine learning approach," *Biomedicines*, vol. 10, no. 8, p. 1897, 2022.

[32] A. Naro, M. G. Maggio, A. Leo, and R. S. Calabro, "Multiplex and multilayer network EEG analyses: A novel strategy in the differential diagnosis of patients with chronic disorders of consciousness," *International Journal of Neural Systems*, vol. 31, no. 02, p. 2050052, 2021.

[33] J. D. Sitt, J.-R. King, I. El Karoui, B. Rohaut, F. Faugeras, A. Gramfort, L. Cohen, M. Sigman, S. Dehaene, and L. Naccache, "Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state," *Brain*, vol. 137, no. 8, pp. 2258–2270, 2014.

[34] D. Engemann, F. Raimondo, J.-R. King, M. Jas, A. Gramfort, S. Dehaene, L. Naccache, and J. Sitt, "Automated measurement and prediction of consciousness in vegetative and minimally conscious patients," in *ICML Workshop on Statistics, Machine Learning and Neuroscience (Stamlins 2015)*, 2015.

[35] Y. Bai, J. He, X. Xia, Y. Wang, Y. Yang, H. Di, X. Li, and U. Ziemann, "Spontaneous transient brain states in EEG source space in disorders of consciousness," *NeuroImage*, vol. 240, p. 118407, 2021.

[36] Y. Höller, A. Thomschewski, J. Bergmann, M. Kronbichler, J. S. Crone, E. V. Schmid, K. Butz, P. Höller, R. Nardone, and E. Trinka, "Connectivity biomarkers can differentiate patients with different levels of consciousness," *Clinical Neurophysiology*, vol. 125, no. 8, pp. 1545–1555, 2014.

[37] U. Lee, S. Ku, G. Noh, S. Baek, B. Choi, and G. A. Mashour, "Disruption of frontal–parietal communication by ketamine, propofol, and sevoflurane," *Anesthesiology*, vol. 118, no. 6, pp. 1264–1275, 2013.

[38] R. Lehembre, M.-A. Bruno, A. Vanhaudenhuyse, C. Chatelle, V. Cologan, Y. Leclercq, A. Soddu, B. Macq, S. Laureys, and Q. Noirhomme, "Resting-state EEG study of comatose patients: a connectivity and frequency analysis to find differences between vegetative and minimally conscious states," *Functional Neurology*, vol. 27, no. 1, p. 41, 2012.

[39] L. E. Ismail and W. Karwowski, "A graph theory-based modeling of functional brain connectivity based on EEG: a systematic review in the context of neuroergonomics," *Ieee Access*, vol. 8, pp. 155 103–155 135, 2020.

[40] C. Duclos, C. Maschke, Y. Mahdid, K. Berkun, J. da Silva Castanheira, V. Tarnal, P. Picton, G. Vanini, G. Golmirzaie, E. Janke *et al.*, "Differential classification of states of consciousness using envelope-and phase-based functional connectivity," *Neuroimage*, vol. 237, p. 118171, 2021.

[41] Z. Hao *et al.*, "Uncovering brain network insights for prognosis in disorders of consciousness: EEG source space analysis and brain dynamics," *IEEE Trans. Neural Sys. and Rehabilitation Eng.*, 2023.

[42] J. Rizkallah *et al.*, "Decreased integration of EEG source-space networks in disorders of consciousness," *NeuroImage: Clinical*, vol. 23, p. 101841, 2019.

[43] D. Nadin *et al.*, "Brain network motif topography may predict emergence from disorders of consciousness: a case series," *Neuroscience of Consciousness*, vol. 2020, no. 1, p. niaa017, 2020.

[44] T. Kustermann *et al.*, "Brain functional connectivity during the first day of coma reflects long-term outcome," *NeuroImage: Clinical*, vol. 27, p. 102295, 2020.

[45] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc Royal Soc. London. Series A: Math. Phy. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, 1998.

[46] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2013.

[47] D. Puri, S. Nalbalwar, A. Nandgaonkar, and A. Wagh, "Alzheimer's disease detection from optimal electroencephalogram channels and tunable Q-wavelet transform," *Indo. Journal of Elec. Engg and Comp. Sci*, vol. 25, no. 3, pp. 1420–1428, 2022.

[48] S. Taran, S. K. Khare, V. Bajaj, and G. Sinha, "Classification of motor-imagery tasks from EEG signals using the rational dilation wavelet transform," in *Modelling and Analysis of Active Biopotential Signals in Healthcare, Volume 2*. IOP Publishing Bristol, UK, 2020, pp. 1–1.

[49] M. Deshmukh, M. Khemchandani, and P. M. Thakur, "Comparative analysis of electroencephalogram (eeg) data gathered from the frontal region with other brain regions affected by attention deficit hyperactivity disorder (adhd)

through multiresolution analysis and machine learning techniques," *Applied Neuropsychology: Child*, pp. 1–15, 2024.

[50] C. M. Sweeney-Reed, S. J. Nasuto, M. F. Vieira, and A. O. Andrade, "Empirical mode decomposition and its extensions applied to EEG analysis: a review," *Advances in Data Science and Adaptive Analysis*, vol. 10, no. 02, p. 1840001, 2018.

[51] J.-R. Yeh, J.-S. Shieh, and N. E. Huang, "Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method," *Advances in adaptive data analysis*, vol. 2, no. 02, pp. 135–156, 2010.

[52] J. Jia, B. Goparaju, J. Song, R. Zhang, and M. B. Westover, "Automated identification of epileptic seizures in EEG signals based on phase space representation and statistical features in the CEEMD domain," *Biomedical Signal Processing and Control*, vol. 38, pp. 148–157, 2017.

[53] V. R. Carvalho, M. F. Moraes, A. P. Braga, and E. M. Mendes, "Evaluating five different adaptive decomposition methods for EEG signal seizure detection and classification," *Biomedical Signal Processing and Control*, vol. 62, p. 102073, 2020.

[54] J. Mathew, N. Sivakumaran, and P. Karthick, "Variational mode decomposition based moment fusion for the detection of seizure types from the scalp EEG measurements," *IEEE Transactions on Instrumentation and Measurement*, 2023.

[55] S. Siuly, S. K. Khare, V. Bajaj, H. Wang, and Y. Zhang, "A computerized method for automatic detection of schizophrenia using EEG signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 11, pp. 2390–2400, 2020.

[56] Z. Aslan, "A novel computer-aided diagnostic system for Alzheimer's diagnosis using variational mode decomposition method," *Circuits, Systems, and Signal Processing*, pp. 1–19, 2023.

[57] S. K. Khare, N. B. Gaikwad, and V. Bajaj, "VHERS: a novel variational mode decomposition and Hilbert transform-based EEG rhythm separation for automatic ADHD detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–10, 2022.

[58] S. K. Khare and V. Bajaj, "An evolutionary optimized variational mode decomposition for emotion recognition," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 2035–2042, 2020.

[59] A. Upadhyay and R. B. Pachori, "Instantaneous voiced/non-voiced detection in speech signals based on variational mode decomposition," *Journal of the Franklin Institute*, vol. 352, no. 7, pp. 2679–2707, 2015.

[60] M. Craig, R. Adapa, I. Pappas, D. Menon, and E. Stamatakis, "Deep graph convolutional neural networks identify frontoparietal control and default mode network contributions to mental imagery manuscript," in *2018 Conference on Cognitive Computational Neuroscience (Philadelphia, PA)*, 2018, pp. 1–8.

[61] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[62] N. Salankar, P. Mishra, and L. Garg, "Emotion recognition from EEG signals using empirical mode decomposition and second-order difference plot," *Biomedical Signal Processing and Control*, vol. 65, p. 102389, 2021.

[63] L. Hu and Z. Zhang, *EEG Signal Processing and Feature Extraction*. Springer, 2019.

[64] T. S. Lorig and G. E. Schwartz, "Factor analysis of the eeg indicates inconsistencies in traditional frequency bands." *Journal of Psychophysiology*, 1989.

[65] Z. Li, L. Zhang, F. Zhang, R. Gu, W. Peng, and L. Hu, "Demystifying signal processing techniques to extract resting-state eeg features for psychologists," *Brain Science Advances*, vol. 6, no. 3, pp. 189–209, 2020.

[66] A. Thul, J. Lechinger, J. Donis, G. Michitsch, G. Pichler, E. F. Kochs, D. Jordan, R. Ilg, and M. Schabus, "EEG entropy measures indicate decrease of cortical information processing in disorders of consciousness," *Clinical Neurophysiology*, vol. 127, no. 2, pp. 1419–1427, 2016.

[67] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical review letters*, vol. 88, no. 17, p. 174102, 2002.

[68] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American journal of physiology-heart and circulatory physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[69] M. Aboy, D. Cuesta-Frau, D. Austin, and P. Mico-Tormos, "Characterization of sample entropy in the context of biomedical signal analysis," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2007, pp. 5942–5945.

[70] J. D. Gibson, "What is the interpretation of spectral entropy?" in *Proceedings of 1994 IEEE International Symposium on Information Theory*. IEEE, 1994, p. 440.

[71] T. Schmierer, T. Li, and Y. Li, "A novel empirical wavelet SODP and spectral entropy based index for assessing the depth of anaesthesia," *Health Information Science and Systems*, vol. 10, no. 1, p. 10, 2022.

[72] S. A. Khoshnevis and R. Sankar, "Applications of higher order statistics in electroencephalography signal processing: A comprehensive survey," *IEEE Reviews in Biomedical Engineering*, vol. 13, pp. 169–183, 2019.

[73] H. Zhao, J. Cao, J. Xie, W.-H. Liao, Y. Lei, H. Cao, Q. Qu, and C. Bowen, "Wearable sensors and features for diagnosis of neurodegenerative diseases: A systematic review," *Digital Health*, vol. 9, p. 20552076231173569, 2023.

[74] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Recognition, 2ed, An Indian Adaptation*. John Wiley & Sons, 2021.

[75] J. S. Clark, P. Kulig, K. Podsiadło, K. Rydzewska, K. Arabski, M. Białecka, K. Safranow, and A. Ciechanowicz, "Empirical investigations into Kruskal-Wallis power studies utilizing Bernstein fits, simulations and medical study datasets," *Scientific Reports*, vol. 13, no. 1, p. 2352, 2023.

[76] M. Staniek and K. Lehnertz, "Symbolic transfer entropy," *Physical review letters*, vol. 100, no. 15, p. 158101, 2008.

[77] T. Bossomaier, L. Barnett, M. Harré, J. T. Lizier, T. Bossomaier, L. Barnett, M. Harré, and J. T. Lizier, *Transfer entropy*. Springer, 2016.

[78] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 67, no. 2, pp. 301–320, 2005.

[79] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation," *Peerj computer science*, vol. 7, p. e623, 2021.

[80] A. Ambat and J. Sahoo, "Anomaly detection and prediction of energy consumption for smart homes using machine learning," *ETRI Journal*, 2024.

[81] Y. Ren, L. Zhang, and P. N. Suganthan, "Ensemble classification and regression-recent developments, applications and future directions," *IEEE Computational intelligence magazine*, vol. 11, no. 1, pp. 41–53, 2016.

[82] J. A. Weaver, A. M. Cogan, K. A. O'brien, P. Hansen, J. T. Giacino, J. Whyte, T. Bender Pape, P. Van Der Wees, and T. Mallinson, "Determining the hierarchy of coma recovery scale-revised rating scale categories and alignment with aspen consensus criteria for patients with brain injury: a rasch analysis," *Journal of neurotrauma*, vol. 39, no. 19-20, pp. 1417–1428, 2022.

[83] A. Anusha and A. Ramakrishnan, "Brain functional connectivity as biomarker for propofol-induced alterations of consciousness," in *2021 43rd Ann. Int. Conf. IEEE EMBC*. IEEE, 2021, pp. 1928–1931.

[84] C. J. Stam, G. Nolte, and A. Daffertshofer, "Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources," *Human Brain Mapping*, vol. 28, no. 11, pp. 1178–1193, 2007.

[85] M. Vinck, R. Oostenveld, M. Van Wingerden, F. Battaglia, and C. M. Pennartz, "An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias," *Neuroimage*, vol. 55, no. 4, pp. 1548–1565, 2011.

[86] L. S. Imperatori, M. Betta, L. Cecchetti, A. Canales-Johnson, E. Ricciardi, F. Siclari, P. Pietrini, S. Chennu, and G. Bernardi, "EEG functional connectivity metrics wPLI and wSMI account for distinct types of brain functional interactions," *Scientific Reports*, vol. 9, no. 1, p. 8894, 2019.

[87] A. Desowska, C. B. Berde, and L. Cornelissen, "Emerging functional connectivity patterns during sevoflurane anaesthesia in the developing human brain," *Brit. Jl Anaesthesia*, vol. 130, no. 2, pp. e381–e390, 2023.

[88] A. Anusha and A. Ramakrishnan, "Midline EEG functional connectivity as biomarker for conscious states in sleep and wakefulness," in *2021 43rd Ann. Int. Conf. IEEE EMBC*. IEEE, 2021, pp. 1924–1927.

[89] B. Schorr, W. Schlee, M. Arndt, and A. Bender, "Coherence in resting-state EEG as a predictor for the recovery from unresponsive wakefulness syndrome," *Journal of Neurology*, vol. 263, pp. 937–953, 2016.

[90] A. Bruns, R. Eckhorn, H. Jokeit, and A. Ebner, "Amplitude envelope correlation detects coupling among incoherent brain signals," *NeuroReport*, vol. 11, no. 7, pp. 1509–1514, 2000.

[91] N. Laskaris, D. Adamos, and A. Bezerianos, "Graph theory for brain signal processing," in *Handbook of Neuroeng.* Springer, 2021, pp. 1–29.

[92] F. Mokhtari, M. I. Akhlaghi, S. L. Simpson, G. Wu, and P. J. Laurienti, "Sliding window correlation analysis: Modulating window shape for dynamic brain connectivity in resting state," *Neuroimage*, vol. 189, pp. 655–666, 2019.

[93] V. Mäki-Marttunen, "Brain dynamic functional connectivity in patients with disorders of consciousness," *BMC Neuroscience*, vol. 15, no. Suppl 1, p. P105, 2014.

[94] C. A. Bareham, N. Roberts, J. Allanson, P. J. Hutchinson, J. D. Pickard, D. K. Menon, and S. Chennu, "Bedside eeg predicts longitudinal behavioural changes in disorders of consciousness," *NeuroImage: Clinical*, vol. 28, p. 102372, 2020.

[95] Y. Guo *et al.*, "Disrupted multi-scale topological organization of directed functional brain networks in patients with disorders of consciousness," *Brain Communications*, vol. 5, no. 2, p. fcad069, 2023.

[96] Y. Li, J. Gao, Y. Yang, Y. Zhuang, Q. Kang, X. Li, M. Tian, H. Lv, and J. He, "Temporal and spatial variability of dynamic microstate brain network

in disorders of consciousness," *CNS Neuroscience & Therapeutics*, vol. 30, no. 2, p. e14641, 2024.

[97] S. Raveendran, R. Kenchaiah, S. Kumar, J. Sahoo, M. Farsana, R. Chowdary Mundlamuri, S. Bansal, V. Binu, A. Ramakrishnan, and S. Ramakrishnan, "Variational mode decomposition-based EEG analysis for the classification of disorders of consciousness," *Frontiers in Neuroscience*, vol. 18, p. 1340528, 2024.

# Curriculum Vitae

Sreelakshmi Raveendran obtained her B.Tech degree in Electronics and Communication Engineering from Government Engineering College, Idukki, Kerala, India in 2013. She received M.Tech degree in Signal Processing from College of Engineering Poonjar, Kerala, India in 2015. She is currently pursuing her Ph.D in the area of EEG signal processing, neuroscience, and machine learning at IIIT Kottayam, Kerala, India. She worked as Academic Associate in the department "IT and Systems Area" in Indian Institute of Management, Kozhikode, India. She also worked as Assistant Professor – Adhoc Faculty, in National Institute of Technology Calicut, Kerala, India from December 11, 2017 to May 9, 2018 and July 10, 2017 to November 30, 2017. Her research interests include resting state EEG, Disorders of consciousness, signal processing, neuroscience, and machine learning.

1. **Bio-data**

   - *Name*: Sreelakshmi Raveendran
   - *Roll No.*: 2019PHD12003
   - *Father's Name*: Raveendranath C R
   - *Date of Birth*: 02/06/1991
   - *Permanent Address*: Chonkara House
     Edanad PO
     Arunapuram Via
     Pala, Kottayam
     Kerala, India
     686574