EMOTION RECOGNITION BASED ON ELECTROENCEPHALOGAPHY (EEG) SIGNALS

A THESIS SUBMITTED TO

THE FACULTY OF ACHITECTURE AND ENGINEERING

OF

EPOKA UNIVERSITY

BY

TEA MALASI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR BACHELOR DEGREE

IN SOFTWARE ENGINEERING

JUNE, 2024

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name, Last name: …………………………..

Signature:

# ABSTRACT

EMOTION RECOGNITION BASED ON ELECTROENCEPHALOGAPHY (EEG) SIGNALS

Malasi, Tea

M.Sc., Department of Computer Engineering

Supervisor: Dr. Florenc Skuka

Emotion recognition has garnered significant attention across various disciplines, with Electroencephalography (EEG) emerging as a promising modality for decoding human emotions. This bachelor thesis investigates the feasibility and efficacy of utilizing EEG signals from the DREAMER dataset for emotion recognition. The study begins with a comprehensive overview of emotion characterization, encompassing both discrete and multi-dimensional perspectives. Additionally, it explores the intricate anatomy of the human brain and outlines the process of EEG data acquisition. Emphasizing EEG's role in emotion recognition, the thesis addresses electrode positioning, artifact mitigation techniques, and the establishment of a robust framework for emotion recognition.

The literature review delves into the methodologies employed in emotion recognition tasks, highlighting machine learning and deep learning techniques. Several datasets, including DEAP, SEED, and MAHNOB HCI-tagging, are discussed, with the experimental focus on the DREAMER dataset. Methodologically, the study details preprocessing steps to enhance EEG data quality, such as bandpass filtering, power spectral density calculation, and baseline-stimulus ratio calculation. It also elaborates on feature extraction techniques in the frequency and time domains and discusses various metrics used to evaluate the accuracy of emotion recognition models.

The results and discussion section presents insights gathered from experiments conducted on the DREAMER dataset, demonstrating the efficacy of the proposed methodologies in emotion recognition tasks. The thesis concludes by suggesting future research directions and highlighting the potential applications of EEG-based emotion recognition systems in diverse domains.

Overall, this thesis advances our understanding of emotion recognition through EEG analysis using the DREAMER dataset and lays the groundwork for practical applications in healthcare, human-computer interaction, and affective computing.

**Keywords:** Electroencephalography (EEG); Machine Learning; Deep Learning; Emotion Recognition; DREAMER dataset.

***Dedicated to my family***

# ACKNOWLEDGEMENTS

My thesis advisor, Dr. Florenc Skuka, has my sincere gratitude for all of his outstanding advice, steadfast support, and ongoing encouragement during the study process. Their knowledge and guidance have been really helpful in developing my thesis.

I also like to thank the Department of Computer Engineering for setting up a setting that supports learning and academic growth. The department's facilities and resources were invaluable for my study, even if I did not directly use them in my thesis.

I want to sincerely thank my family and friends for their unwavering support, love, and understanding throughout this journey. Their unwavering support has been as a constant source of inspiration and strength.

Finally, I would like to acknowledge my classmates for the support they provided me during the entire informative course as well as their cooperation and mere presence, which enhanced my learning process.

Last but not the least, I would like to extend my deepest gratitude to all of those individuals who in one way or the other have assisted in the completion of the graduation project. I would like to strongly appreciate your support and assistance you have provided during this process.

# TABLE OF CONTENTS

[ABSTRACT 3](#_Toc168610182)

[ACKNOWLEDGEMENTS 6](#_Toc168610183)

[TABLE OF CONTENTS 7](#_Toc168610184)

[LIST OF TABLES 10](#_Toc168610185)

[LIST OF FIGURES 14](#_Toc168610186)

[LIST OF ABBREVIATIONS 16](#_Toc168610187)

[CHAPTER 1 1](#_Toc168610188)

[INTRODUCTION 1](#_Toc168610189)

[1.1. Emotions Characterization 1](#_Toc168610190)

[1.1.1. Discrete Emotions 2](#_Toc168610191)

[1.1.2. Multi-dimensional Emotions 3](#_Toc168610192)

[1.2. Human Brain Anatomy 4](#_Toc168610193)

[1.3. Data Acquisition 5](#_Toc168610194)

[1.4. EEG In Emotion Recognition 6](#_Toc168610195)

[1.4.1. Electrodes Position 8](#_Toc168610196)

[1.4.2. Artifacts 9](#_Toc168610197)

[1.4.3. EEG And Human Brain Connection 10](#_Toc168610198)

[1.4.4. The Framework of Emotion Recognition 11](#_Toc168610199)

[1.5. Aim and Novelty of the Study 11](#_Toc168610200)

[CHAPTER 2 13](#_Toc168610201)

[LITERATURE REVIEW 13](#_Toc168610202)

[2.1. Methods Used for Emotion Recognition 13](#_Toc168610203)

[2.2.1. Machine Learning Methods and Results 15](#_Toc168610204)

[2.2.2. Deep Learning Methods and Results 20](#_Toc168610205)

[2.2. Datasets 25](#_Toc168610206)

[2.2.1. DEAP *---* Database for emotion analysis using physiological signals 25](#_Toc168610207)

[2.2.2. SEED *---* SJTU emotion EEG dataset 26](#_Toc168610208)

[2.2.3. DREAMER 27](#_Toc168610209)

[2.2.4. MAHNOB HCI- tagging database 28](#_Toc168610210)

[2.3. Metrics 29](#_Toc168610211)

[CHAPTER 3 32](#_Toc168610212)

[METHODOLOGY 32](#_Toc168610213)

[3.1. Datasets Used in This Research 32](#_Toc168610214)

[3.1.1. Dreamer Dataset 32](#_Toc168610215)

[3.2. Data Preprocessing 35](#_Toc168610216)

[3.2.1. Bandpass Filtering (BPF) 35](#_Toc168610217)

[3.2.2. Power Spectral Density Calculation (PSD) 36](#_Toc168610218)

[3.2.3. Baseline-Stimulus Ratio Calculation 36](#_Toc168610219)

[3.2.4. Scaling 36](#_Toc168610220)

[3.2.5. Feature Normalization 37](#_Toc168610221)

[3.3. Feature Extraction 37](#_Toc168610222)

[3.3.1. Frequency Domain 38](#_Toc168610223)

[3.3.2. Time Domain and Statistical 38](#_Toc168610224)

[3.4. Machine Learning Methods used in this Research 40](#_Toc168610225)

[3.4.1. Support Vector Machine (SVM) 41](#_Toc168610226)

[3.4.2. Random Forest (RF) 41](#_Toc168610227)

[3.4.3. Decision Tree (DT) 41](#_Toc168610228)

[3.4.4. K-Nearest Neighbor (K-NN) 42](#_Toc168610229)

[3.5. Deep Learning Methods used in this Study 42](#_Toc168610230)

[3.5.1. Conventual Neural Networks (CNN) 43](#_Toc168610231)

[3.5.2. Recurrent Neural Network 43](#_Toc168610232)

[3.5.3. Deep Neural Network 44](#_Toc168610233)

[3.6. Metrics used in this Study 44](#_Toc168610234)

[CHAPTER 4 45](#_Toc168610235)

[RESULTS AND DISCUSSION 45](#_Toc168610236)

[4.1. Frequency Domain Results on Machine Learning Algorithms 45](#_Toc168610237)

[4.1.1. Binary Classification 47](#_Toc168610238)

[4.1.2. Ternary Classification 53](#_Toc168610239)

[4.2. Time Domain Results on Machine Learning Algorithms 59](#_Toc168610240)

[4.2.1. Binary Classification 60](#_Toc168610241)

[4.2.2. Ternary Classification 66](#_Toc168610242)

[4.3. Frequency Domain Results on Deep Learning Algorithms 72](#_Toc168610243)

[4.3.1. Binary Classification 73](#_Toc168610244)

[4.3.2. Ternary Classification 74](#_Toc168610245)

[4.4. Time Domain Results on Deep Learning Algorithms 75](#_Toc168610246)

[4.4.1. Binary Classification 75](#_Toc168610247)

[4.4.2. Ternary Classification 76](#_Toc168610248)

[CONCLUSION 78](#_Toc168610249)

[REFERENCES 79](#_Toc168610250)

# LIST OF TABLES

TABLES

[Table 1. Association between brain areas and EEG frequency bands used to measure activity levels and mental state. 10](#_Toc168607355)

[Table 2. Confusion matrix sample. 31](#_Toc168607356)

[Table 3. jth column of EEG recordings corresponds to the following EEG positions. 34](#_Toc168607357)

[Table 4. Table that shows the best parameters of RF. 47](#_Toc168607358)

[Table 5. Table that shows the best parameters of SVM. 47](#_Toc168607359)

[Table 6.Table that shows the best parameters of DT. 47](#_Toc168607360)

[Table 7. Table with results of valance for frequency domain, binary classification using machine learning methods. 48](#_Toc168607361)

[Table 8. Table that shows the best parameters of RF. 49](#_Toc168607362)

[Table 9. Table that shows the best parameters of SVM. 49](#_Toc168607363)

[Table 10.Table that shows the best parameters of DT. 49](#_Toc168607364)

[Table 11. Table with results of arousal for frequency domain, binary classification using machine learning methods. 50](#_Toc168607365)

[Table 12. Table that shows the best parameters of RF. 51](#_Toc168607366)

[Table 13. Table that shows the best parameters of SVM. 51](#_Toc168607367)

[Table 14.Table that shows the best parameters of DT. 51](#_Toc168607368)

[Table 15. Table with results of dominance for frequency domain, binary classification using machine learning methods. 52](#_Toc168607369)

[Table 16. Table that shows the best parameters of RF. 53](#_Toc168607370)

[Table 17. Table that shows the best parameters of SVM. 53](#_Toc168607371)

[Table 18.Table that shows the best parameters of DT. 53](#_Toc168607372)

[Table 19. Table with results of valence for frequency domain, ternary classification using machine learning methods. 54](#_Toc168607373)

[Table 20. Table that shows the best parameters of RF. 55](#_Toc168607374)

[Table 21. Table that shows the best parameters of SVM. 55](#_Toc168607375)

[Table 22.Table that shows the best parameters of DT. 55](#_Toc168607376)

[Table 23. Table with results of arousal for frequency domain, ternary classification using machine learning methods. 56](#_Toc168607377)

[Table 24. Table that shows the best parameters of RF. 57](#_Toc168607378)

[Table 25. Table that shows the best parameters of SVM. 57](#_Toc168607379)

[Table 26.Table that shows the best parameters of DT. 57](#_Toc168607380)

[Table 27. Table with results of dominance for frequency domain, ternary classification using machine learning methods. 58](#_Toc168607381)

[Table 28. Table that shows the best parameters of RF. 60](#_Toc168607382)

[Table 29. Table that shows the best parameters of SVM. 60](#_Toc168607383)

[Table 30.Table that shows the best parameters of DT. 61](#_Toc168607384)

[Table 31.Table with results of valance for time domain, binary classification using machine learning methods. 61](#_Toc168607385)

[Table 32. Table that shows the best parameters of RF. 62](#_Toc168607386)

[Table 33. Table that shows the best parameters of SVM. 62](#_Toc168607387)

[Table 34.Table that shows the best parameters of DT. 63](#_Toc168607388)

[Table 35. Table with results of arousal for time domain, binary classification using machine learning methods. 63](#_Toc168607389)

[Table 36. Table that shows the best parameters of RF. 64](#_Toc168607390)

[Table 37. Table that shows the best parameters of SVM. 64](#_Toc168607391)

[Table 38.Table that shows the best parameters of DT. 65](#_Toc168607392)

[Table 39. Table with results of dominance for time domain, binary classification using machine learning methods. 65](#_Toc168607393)

[Table 40. Table that shows the best parameters of RF. 66](#_Toc168607394)

[Table 41. Table that shows the best parameters of SVM. 66](#_Toc168607395)

[Table 42.Table that shows the best parameters of DT. 67](#_Toc168607396)

[Table 43. Table with results of valence for time domain, ternary classification using machine learning methods. 67](#_Toc168607397)

[Table 44. Table that shows the best parameters of RF. 68](#_Toc168607398)

[Table 45. Table that shows the best parameters of SVM. 68](#_Toc168607399)

[Table 46.Table that shows the best parameters of DT. 69](#_Toc168607400)

[Table 47. Table with results of arousal for time domain, ternary classification using machine learning methods. 69](#_Toc168607401)

[Table 48. Table that shows the best parameters of RF. 70](#_Toc168607402)

[Table 49. Table that shows the best parameters of SVM. 70](#_Toc168607403)

[Table 50.Table that shows the best parameters of DT. 71](#_Toc168607404)

[Table 51. Table with results of dominance for time domain, ternary classification using machine learning methods. 71](#_Toc168607405)

[Table 52. Table with results of valence for frequency domain, binary classification using deep learning methods. 73](#_Toc168607406)

[Table 53. Table with results of arousal for frequency domain, binary classification using deep learning methods. 73](#_Toc168607407)

[Table 54. Table with results of dominance for frequency domain, binary classification using deep learning methods. 74](#_Toc168607408)

[Table 55. Table with results of valence for frequency domain, ternary classification using deep learning methods. 74](#_Toc168607409)

[Table 56. Table with results of arousal for frequency domain, ternary classification using deep learning methods. 74](#_Toc168607410)

[Table 57. Table with results of dominance for frequency domain, ternary classification using deep learning methods. 75](#_Toc168607411)

[Table 58. Table with results of valence for time domain, binary classification using deep learning methods. 76](#_Toc168607412)

[Table 59. Table with results of arousal for time domain, binary classification using deep learning methods. 76](#_Toc168607413)

[Table 60. Table with results of dominance for frequency domain, binary classification using deep learning methods. 76](#_Toc168607414)

[Table 61. Table with results of valence for time domain, ternary classification using deep learning methods. 77](#_Toc168607415)

[Table 62. Table with results of arousal for time domain, ternary classification using deep learning methods. 77](#_Toc168607416)

[Table 63. Table with results of dominance for time domain, ternary classification using deep learning methods. 77](#_Toc168607417)

# LIST OF FIGURES

FIGURES

[Figure 1. Robert Plutchik’s wheel of emotions.(3) 2](#_Toc168912859)

[Figure 2. 2D valence and arousal Paradigms.(1) 3](#_Toc168912860)

[Figure 3. 3D (Valence-Arousal-Dominance) model of emotions.(1) 4](#_Toc168912861)

[Figure 4. The Frontal, Parietal, Temporal, and Occipital lobes of the brain.(5) 5](#_Toc168912862)

[Figure 5. EEG cap visualization.(8) 7](#_Toc168912863)

[Figure 6. Brain wave samples with frequencies.(1) 7](#_Toc168912864)

[Figure 7.Top view of the skull where electrodes are put according to global 10/20 8](#_Toc168912865)

[Figure 8.Removal of Artifacts from EEG Signals. (7) 9](#_Toc168912866)

[Figure 9.The General framework of emotion recognition using AI.(1) 11](#_Toc168912867)

[Figure 10. Methods for classifying in deep learning and machine learning. 14](#_Toc168912868)

[Figure 11.Support Vector Machine algorithm.(9) 16](#_Toc168912869)

[Figure 12.Random Forest algorithm.(11) 17](#_Toc168912870)

[Figure 13. Decision Tree algorithm.(13) 18](#_Toc168912871)

[Figure 14. K Nearest Neighbor algorithm.(14) 19](#_Toc168912872)

[Figure 15. Convolutional Neural Network layout.(16) 21](#_Toc168912873)

[Figure 16. Deep Belief Network feed-forward architecture layout.(17) 22](#_Toc168912874)

[Figure 17. Deep Neural Network feed-forward architecture layout.(18) 23](#_Toc168912875)

[Figure 18. Probabilistic Neural Network architecture layout.(20) 24](#_Toc168912876)

[Figure 19. Video clip valence and arousal values in the DEAP database. Videos that are included are displayed in green.(21) 26](#_Toc168912877)

[Figure 20. KNN performance for different values of k where the best one is k=3. 48](#_Toc168912878)

[Figure 21. KNN performance for different values of k where the best one is k=3. 50](#_Toc168912879)

[Figure 22. KNN performance for different values of k where the best one is k=2. 52](#_Toc168912880)

[Figure 23. KNN performance for different values of k where the best one is k=2. 54](#_Toc168912881)

[Figure 24. KNN performance for different values of k where the best one is k=2. 56](#_Toc168912882)

[Figure 25. KNN performance for different values of k where the best one is k=3. 58](#_Toc168912883)

[Figure 26. KNN performance for different values of k where the best one is k=2. 61](#_Toc168912884)

[Figure 27. KNN performance for different values of k where the best one is k=2. 63](#_Toc168912885)

[Figure 28. KNN performance for different values of k where the best one is k=2. 65](#_Toc168912886)

[Figure 29. KNN performance for different values of k where the best one is k=2. 67](#_Toc168912887)

[Figure 30. KNN performance for different values of k where the best one is k=2. 69](#_Toc168912888)

[Figure 31. KNN performance for different values of k where the best one is k=3. 71](#_Toc168912889)

# LIST OF ABBREVIATIONS

|  |  |
| --- | --- |
|  |  |
| **EEG**  **SVM**  **RF**  **DT**  **K-NN**  **ANN**  **MLP**  **CNN**  **DBN**  **RNN** | Electroencephalography signals  Support Vector Machine  Random Forest  Decision Tree  K-Nearest Neighbor  Artificial Neural Networks  Multi-layer Perception  Convolutional Neural Networks  Deep Belief Networks  Recurrent Neural Networks |
| **DNN**  **BDAE**  **PNN**  **DEAP**  **ECG**  **BPF**  **PSD** | Deep Neural Networks  Bimodal Deep Autoencoder  Probabilistic Neural Network  Database for Emotion Analysis using Physiological signals  Electrocardiographic signals  Band Pass Filtering  Power Spectral Density |

# CHAPTER 1

# INTRODUCTION

Human emotions form the core of our existence, shaping our perceptions, decisions, reactions, feelings, and interactions with the world. Knowing that they are a crucial element of us, that sometimes we can control and sometimes not, makes their study not only a journey into the depths of our psyche but also a crucial exploration for unraveling the complexities of human behavior and fostering deeper understanding within our society. Recognizing emotions solely by our brain signals is an important step into improving what we call human-computer interaction, but an even bigger step in the applications of healthcare.

## Emotions Characterization

In the realm of emotion modeling, we can see have two types of emotions models: discrete emotion models and multi-dimensional emotion models. Discrete emotion models categorize emotions into distinct categories like happiness, sadness, and anger. This model offers a simplified and intuitive structure for comprehending emotional experiences. In contrast, multi-dimensional emotion models acknowledge the intricate nature of emotions. They consider factors such as valence, arousal, and appraisal dimensions, in order to provide a more sophisticated understanding of emotional states and their interrelationships. (1)

### Discrete Emotions

Psychologist Robert Plutchik made a significant classification of emotions in 1980 that is still used in discrete emotion models. The wheel of emotions motive illustrates Plutchik's eight main categories of emotions: anger, surprise, fear, joy, sadness, trust, disgust, and anticipation (Figure 1) (2). The visual representation shows the relationships between different emotions and the outcomes of their combinations. It's notable that emotions intensify as they gravitate towards the center. Rather than offering quantitative evaluations of emotions, these discrete emotion models offer descriptive insights into emotions (1). They help you understand the diversity and complexity of emotional experiences without exploring numerical measurements.

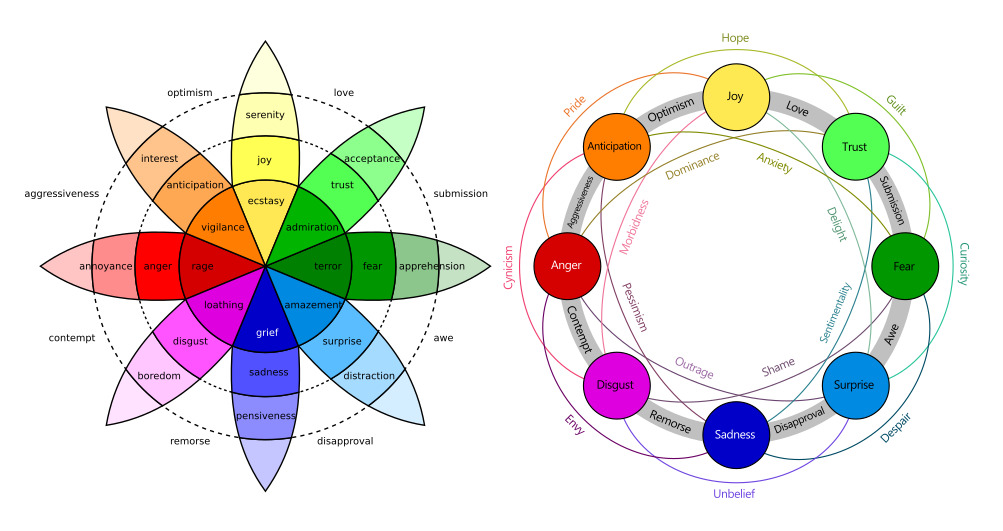


Figure 1. Robert Plutchik’s wheel of emotions.(3)

### Multi-dimensional Emotions

Two different kinds of emotional spaces exist in multi-dimensional space: two-dimensional (2D) and three-dimensional (3D) spaces. Emotions in 2D space can be classified according to their valence and arousal. The degree of either happy or negative emotions following emotional perception is measured by valence. Arousal identifies passive or active emotional states and quantifies the intensity of emotions ranging from low to high. Russell presented a well-known 2D space model. (Figure 2). Further elaborating on this, Mehrabian and Russell created a 3D emotion model (Figure 3), incorporating the VAD (valence, arousal, and dominance) dimensions (1). The third dominating axis in this model, which goes from passive to dominant, represents human control or power in a certain emotional state.

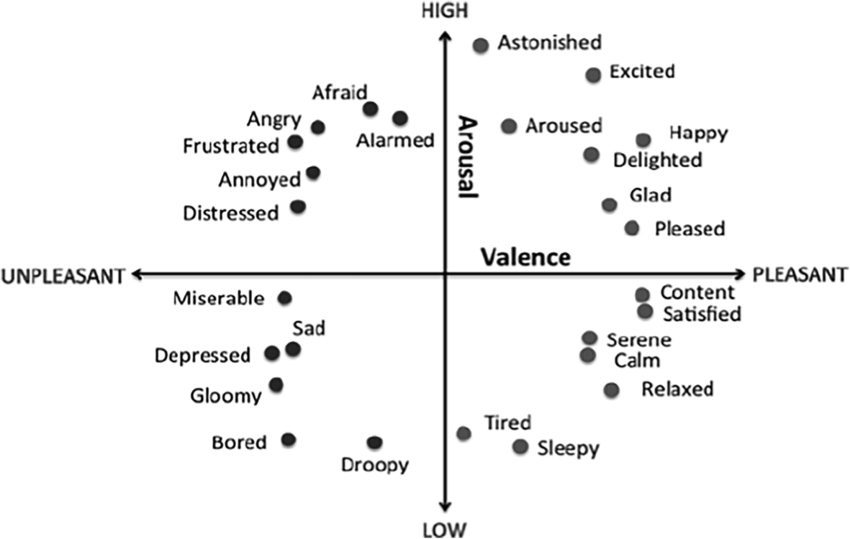


Figure 2. 2D valence and arousal Paradigms.(1)

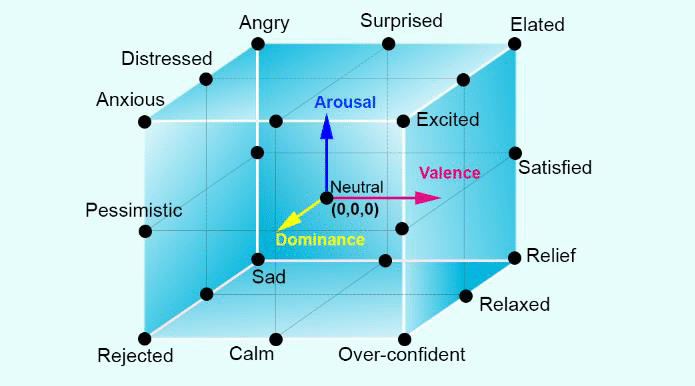


Figure 3. 3D (Valence-Arousal-Dominance) model of emotions.(1)

## Human Brain Anatomy

Understanding the human brain anatomy is a crucial step for ensuring the collection of reliable EEG data. The cerebrum, the cerebellum, and the brainstem are the three main components of the brain(4). Out of the three, the cerebrum is the largest and is separated into the left and right hemispheres (Figure 4). The corpus callosum, a bundle of fibers that connects these hemispheres, is responsible for sending communications between them.

Since the left and right hemispheres of the brain regulate the corresponding sides of the body, a stroke on the right side of the brain could cause paralysis or weakness in the left arm or leg. In terms of function, the right hemisphere is in charge of creativity, spatial aptitude, artistic and musical abilities, and the left hemisphere is in charge of speaking, comprehension, mathematics, and writing.

It is important to note that 92% of people have dominant left hemispheres when it comes to language and hand use. Each cerebral hemisphere is made up of four distinct lobes: the frontal, parietal, temporal, and occipital (Figure 8), which play essential roles contributing to the brain’s intricate functionality.

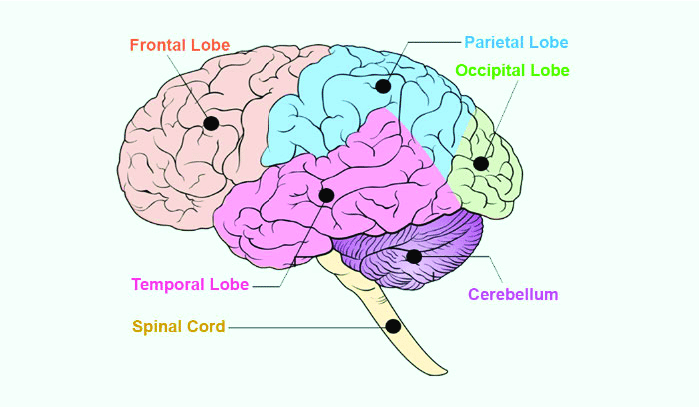


Figure 4. The Frontal, Parietal, Temporal, and Occipital lobes of the brain.(5)

## Data Acquisition

The procedure of obtaining EEG data has been greatly expedited by recent developments, which may lead to increased viability for doing research in less restricted settings. The Mobita 32-channel wireless EEG system, Brain Vision LLC, Emotiv Epoc headset, Biosemi ActiveTwo, and Neuroscan EEG module are a few examples of traditional EEG data collecting devices (4).

Unlike most other systems, the ActiveTwo system records data without a reference, which means that manual subtraction must be done later on in the processing process. It also has the capacity to receive external event code triggers through an external integration box, which is a crucial part of the ActiveTwo system, using a single 16-bit interface that consists of two 8-bit parallel ports. Before any data is transferred to the PC, this integration box directly interfaces with the acquisition devices (6).

The EPOC headset is equipped with 14 data electrodes positioned at predetermined 10–20 electrode locations: AF3, AF4, F7, F3, F4, F8, FC5, FC6, T7, T8, P7, P8, O1, and O2. These electrodes serve various purposes, including data collection, reference and common-mode sensing. Additionally, they allow the users to select their preferred reference locations, utilizing either P3 and P4 or the left/right mastoids, and to place non-conductive rubber pads at alternative locations for customization.

Moreover, the headset incorporates two accelerometers for monitoring head movements, and is cost-effective. However, the EPOC system exhibits limited adaptability to accommodate varying head sizes, posing challenges for users with non-standard head dimensions. It also exclusively facilitates wireless communication between the headset and a USB data collection device via a proprietary 2.4 GHz protocol(6). These considerations are important for researchers evaluating the suitability of the EPOC headset for their EEG data collection needs.

## EEG In Emotion Recognition

With the use of electrodes, EEG is the most widely used brain-imaging technique for assessing electrical activity in the brain. The majority of EEG acquisition hardware is in the form of a cap or headset with different connected and wireless electrodes or sensors attached (Figure 5). In order to improve contact, these electrodes are designed to be applied to the surface of the skull or even the cortex using conductive gel and abrasion. Different electrode types obtain EEG data in different ways. The electrodes listed below are frequently used to capture EEG data (7):

* Disposable Electrodes (Gelled/Pre-Gelled)
* Reusable Electrodes
* Electrodes Caps (Headbands)
* Saline-Based Electrodes
* Needle Electrodes

These electrodes are suitable, convenient, and cost-effective to varying degrees and fulfill a variety of functions. Five major electrical patterns or waves, alpha (α), theta (θ), beta (β), delta (δ), and gamma (γ), can be produced by the human brain (Figure 6). These waves combine to form an EEG signal, which allows us to identify the emotions.

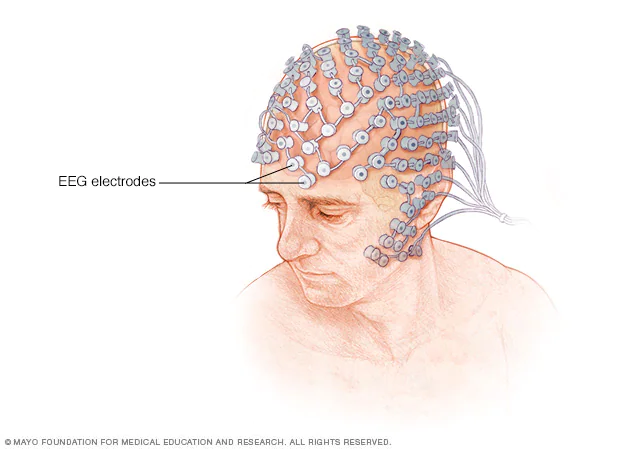


Figure 5. EEG cap visualization.(8)

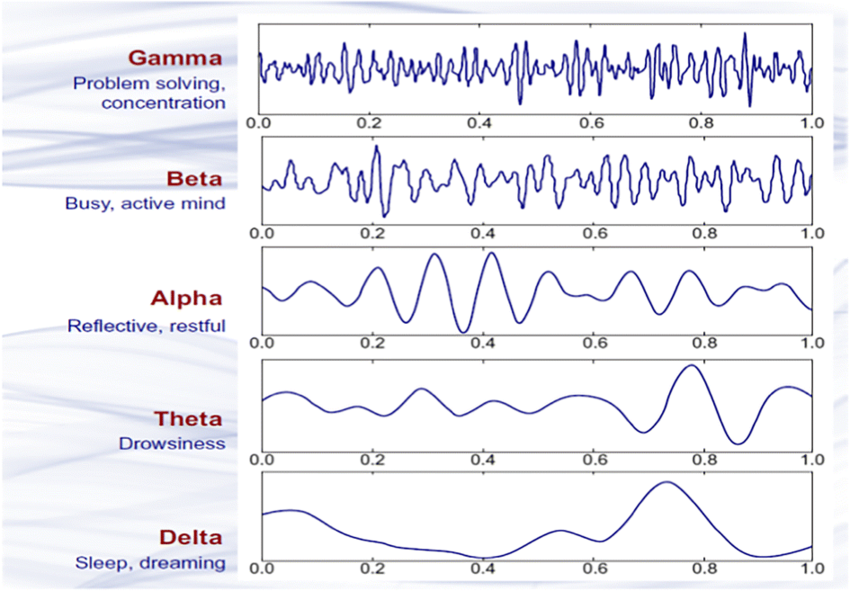


Figure 6. Brain wave samples with frequencies.(1)

### Electrodes Position

The electrodes on the surface of the skull are crucial in determining the EEG signal's amplitude. The International 10/20 electrode placement scheme is commonly used for emotion recognition, as shown in Figure 7 (4). The symbol "10/20" designates a 10% or 20% gap between neighboring electrodes positioned on the scalp (1). Even-numbered electrodes are positioned on the right side of the head, while odd-numbered electrodes are positioned on the left, in accordance with the normal numbering scheme.

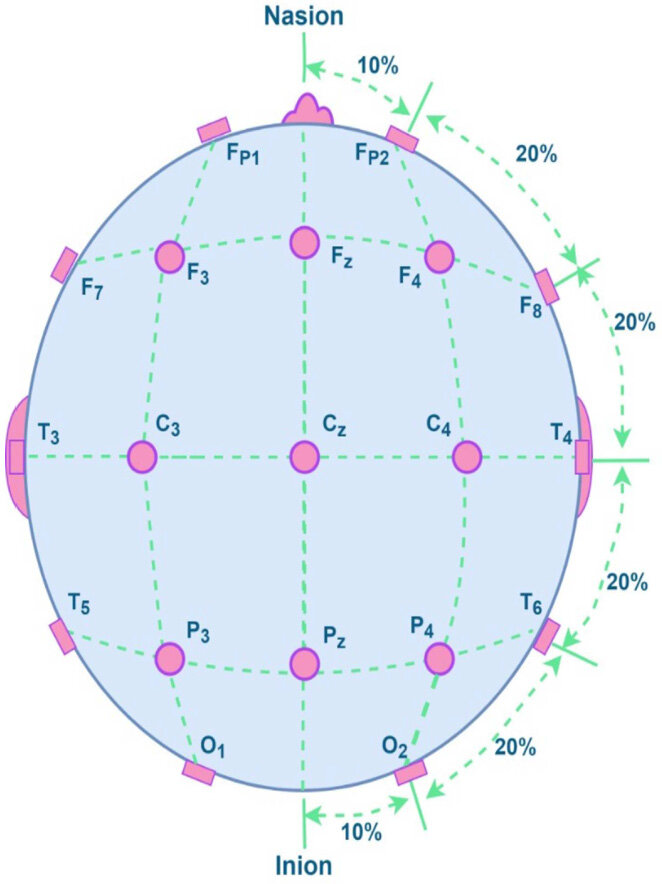


Figure 7.Top view of the skull where electrodes are put according to global 10/20

electrode positioning standard. (4)

### Artifacts

Unwanted signals caused by noise sources are called artifacts. These are caused by interference with EEG signal readings rather than brain activity itself, which makes it challenging to accurately analyze brain activity. To assure the data's dependability, it is necessary to remove such noise from raw EEG signals. Different kinds of artifacts exist.

The system's impedance is one of the primary causes of artifacts, and sampling frequency artifact is another. Impedance artifacts result from variations or anomalies in the impedance of the system, whereas sampling frequency abnormalities are usually caused by ground loop problems at frequencies between 50 and 60 Hz. Using efficient artifact removal algorithms becomes crucial when one realizes how important artifacts are and how they might skew EEG records (Figure 8) (7).



Figure 8.Removal of Artifacts from EEG Signals. (7)

* (A): The raw EEG signal having large artifacts.
* (B): The averaged imaging artifact.
* (C): The result of subtracting the averaged imaging artifact in B from the EEG in A, followed by down-sampling and showing Pulse artifact.
* (D): The averaged pulse artifact from trace C (not to scale).
* (E): Result of subtracting the averaged pulse artifact in D from the EEG in C.
* (F): The EEG from the same subject, recorded outside the scanner, i.e., free of imaging and pulse artifacts.

### EEG And Human Brain Connection

There are associations between distinct mental states and brain regions with each of the five subbands of the EEG signals (Table 1). The representation of each signal, as well as the matching brain areas and intensity of each signal, can be better understood by looking at Figures 6, 7, and 1.(4)

Table 1. Association between brain areas and EEG frequency bands used to measure activity levels and mental state.

|  |  |  |  |
| --- | --- | --- | --- |
| EEG Rhythms | Frequency | Location on brain | Activity |
| Delta (δ) | (0.1 - 4) Hz | Frontal | Sleep, Dream, Babies |
| Theta (θ) | (4 - 8) Hz | Midline, Temporal | Imaginary, Drowsiness, Reflective |
| Alpha (α) | (8 - 13) Hz | Frontal, Occipital | Calm, Relaxed, Eyes closing |
| Beta (β) | (13 - 30) Hz | Frontal, distributed on both sides | Thinking, start to alert, Anxious |
| Gamma (γ) | (>30) Hz | Frontal, Central, Somatosensory cortex | Problem solving, Concertation, Agitation |

### The Framework of Emotion Recognition

The manner in which emotions are elicited holds significance in emotion recognition systems. While some argue that video clips are most effective stimuli for triggering human emotions, others contend that music or memories serve as superior stimuli. However, it is widely acknowledged that the intensity of the stimulation correlates with the richness of the database. By using robust and potent stimuli, enhances the likelihood of achieving superior results and increases the accuracy in emotion recognition.

An AI-based emotion detection system's overall workflow usually begins with feature extraction, which is the process of extracting pertinent information from unprocessed data. To find the most informative features, feature selection is then carried out. After that, a classification algorithm uses these chosen attributes to classify or forecast results. Ultimately, the system's performance is assessed to make sure it is successful in doing the intended goal and to offer insightful information for future developments (Figure 9).

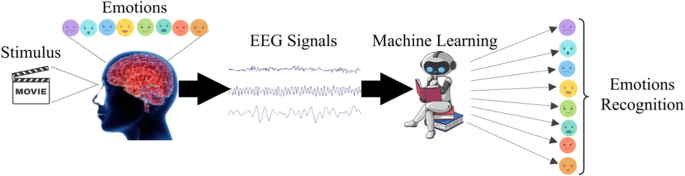


Figure 9.The General framework of emotion recognition using AI.(1)

## Aim and Novelty of the Study

The goal of this study was to explore the potential of Electroencephalography (EEG) signals in recognizing and interpreting human emotions based on valence, arousal and dominance. The study focuses on understanding the intricate relationship between EEG signals and emotional states, utilizing advanced data processing and machine learning techniques.

In order to create a strong framework for emotion recognition based on EEG signals, the research will examine the effectiveness of several approaches, such as deep learning models and conventional machine learning algorithms. By tackling important issues including electrode placement, artifact removal, and feature extraction, the study also aims to provide fresh perspectives into the subject. The ultimate goal is to improve our knowledge of how to recognize emotions in EEG data and open up new avenues for useful applications in the human-computer interface, affective computing, and healthcare domains.

# CHAPTER 2

# LITERATURE REVIEW

## Methods Used for Emotion Recognition

Since the EEG signal has a very little amplitude, it is exceedingly difficult to discern emotion from it. Nonetheless, a lot of scientists have made an effort to solve this problem by utilizing cutting-edge approaches like deep learning and machine learning. There are two primary methods for the whole EEG-based emotion recognition system:

* Deep Learning (DL) based;
* Machine Learning (ML) based.

Understanding the difference between supervised and unsupervised learning techniques is essential for both machine learning and deep learning courses. The key distinction between classification and clustering approaches is that supervised learning uses labeled data for training, whereas unsupervised learning tasks usually entail grouping data without labels.

Multiple-layer neural networks are used in deep learning (DL) techniques to automatically extract features from unprocessed EEG data. Convolutional neural networks (CNN), deep belief networks (DBN), recurrent neural networks (RNN), bimodal deep auto encoders (BDAE), deep neural networks (DNN), voting ensembles (VEn), probabilistic neural networks (PNN), and so on are classifiers found in several deep learning-based systems.

On the other hand, algorithms are used in machine learning (ML) techniques to either cluster data into groups without predefined labels (unsupervised learning) or discover patterns from labeled data (supervised learning). Classifiers like Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Artificial Neural Network (ANN), Multi-Layer Perceptron (MLP), k-Nearest Neighbor (KNN), and others are used in machine learning-based systems (Figure 10)(4). These diverse methodologies offer researchers a range of tools to explore and develop effective emotion recognition systems from EEG signals.

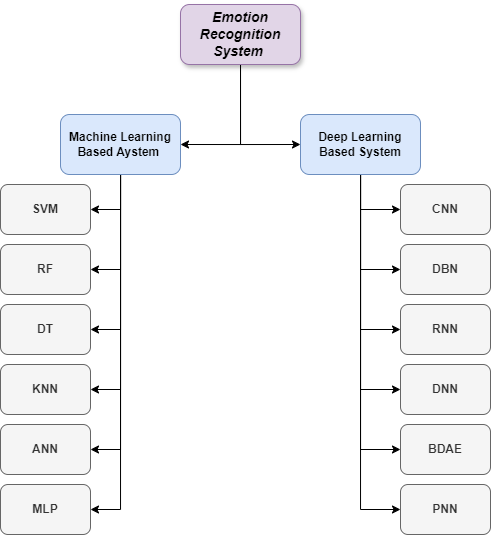


Figure 10. Methods for classifying in deep learning and machine learning.

### Machine Learning Methods and Results

#### Support Vector Machine (SVM)

Support Vector Machine (SVM) works by finding a hyperplane within the feature space that effectively separates data points into two distinct classes. This hyperplane, which can be visualized as a flat surface, serves as the boundary between the two classes. In simpler terms, if the dataset has two features, the hyperplane is represented as a line; for datasets with three features, it becomes a plane, and so forth. To optimize the training of machine learning models, the selection of the best hyperplane is pivotal, and it is achieved by maximizing the distance between each class(1). (Figure 11)

A key component of SVM is the kernel function, which facilitates the transformation of low-dimensional data into higher-dimensional spaces. This transformation enhances the accuracy of data classification while simultaneously reducing computational complexity. Furthermore, it contributes to the generalization capability of the classifier. It is worth mentioning that SVM is a supervised learning algorithm used mainly for classification tasks.

A framework called HAF-HOC was introduced by Petrantonakis et al. The retrieved characteristics were used as SVM's input. According to their study, the accuracy rate was 85%. Furthermore, characteristics collected from EEG and ocular signals were classified using linear SVM into three emotion states (positive, neutral, and negative), resulting in an average identification rate of 91.49% in another study. In the area of EEG-based emotion recognition automatically, SVM showed exceptional accuracy of 99.82% (1).

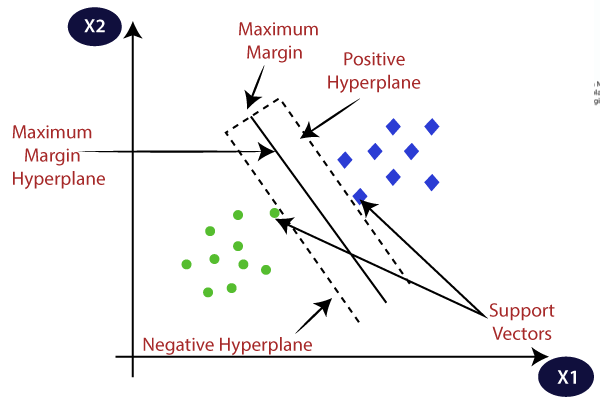


Figure 11.Support Vector Machine algorithm.(9)

#### Random Forest (RF)

The random forest algorithm is widely embraced in the realm of supervised machine learning, thanks to its adaptability in addressing both regression and classification tasks. It belongs to the family of ensemble methods, where predictions from multiple decision trees are combined to enhance accuracy. (Figure 12)

In the domain of analyzing survival data, the random forest method demonstrates commendable predictive prowess. However, its implementation can pose challenges due to its intensive computational requirements and dependence on various input parameters. Furthermore, its "black box" nature presents challenges, as it lacks easily interpretable parameters that elucidate the influence of predictors on outcomes.

Despite these challenges, there are effective strategies to navigate them. By conducting thorough testing of the random forest model, researchers can generate partial effect plots, offering valuable insights into predictor impacts. Nevertheless, when dealing with extensive datasets and numerous predictors, selecting variables for plotting may present challenges. Additionally, attempting to hold non-plotted variables constant could potentially constrain the flexibility of the algorithm.

In essence, while the random forest algorithm boasts robust predictive capabilities, its interpretation and practical application require careful consideration and adjustments to extract meaningful insights from the analysis.(10)

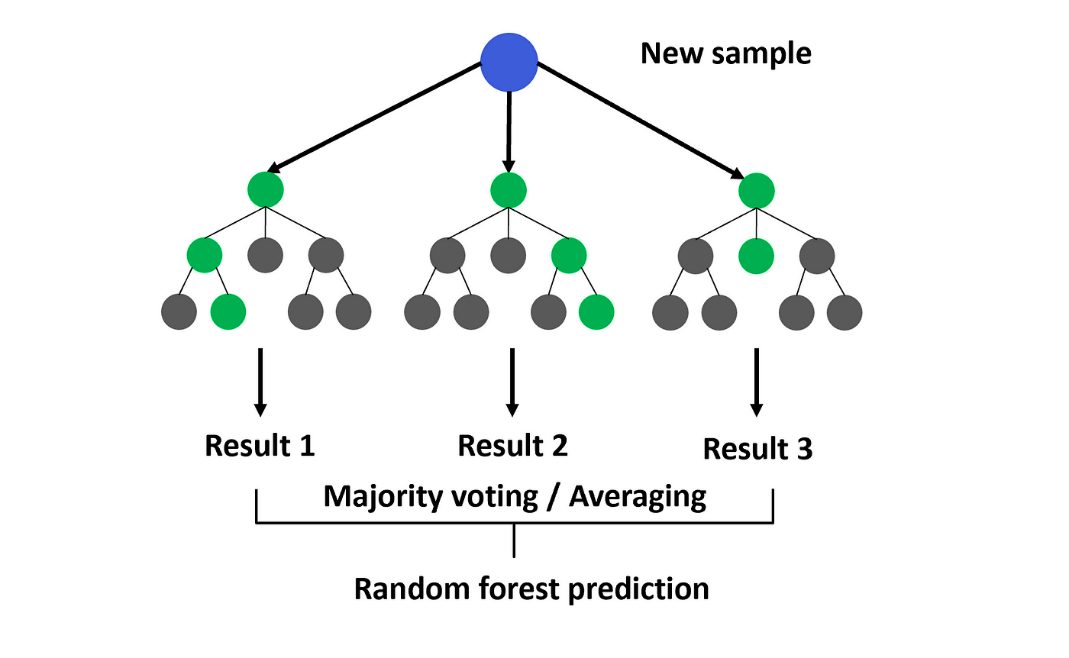


Figure 12.Random Forest algorithm.(11)

#### Decision Tree (DT)

A decision tree structure closely resembles that of a regular tree, comprising a root node, branches, and leaf nodes. At each internal node, an attribute is tested, with the outcome of the test determining the branch direction, and the final class label being assigned to a leaf node (Figure 13).

The root node serves as the parent of all nodes, positioned at the top of the tree. Each node in the decision tree represents a feature or attribute, while the branches indicate decisions or rules. Finally, the leaf nodes represent outcomes, which can be either categorical or continuous values.

Decision trees emulate human-level thinking, making them intuitive for data interpretation. The goal is to construct a tree that encapsulates the entire dataset, processing a single outcome at each leaf node.

One of the key advantages of decision trees is their transparency. They explicitly illustrate all possible alternatives and follow each alternative to its conclusion in a single view, facilitating easy comparison among different options. This transparency is particularly advantageous for understanding the decision-making process and interpreting the results.(12)

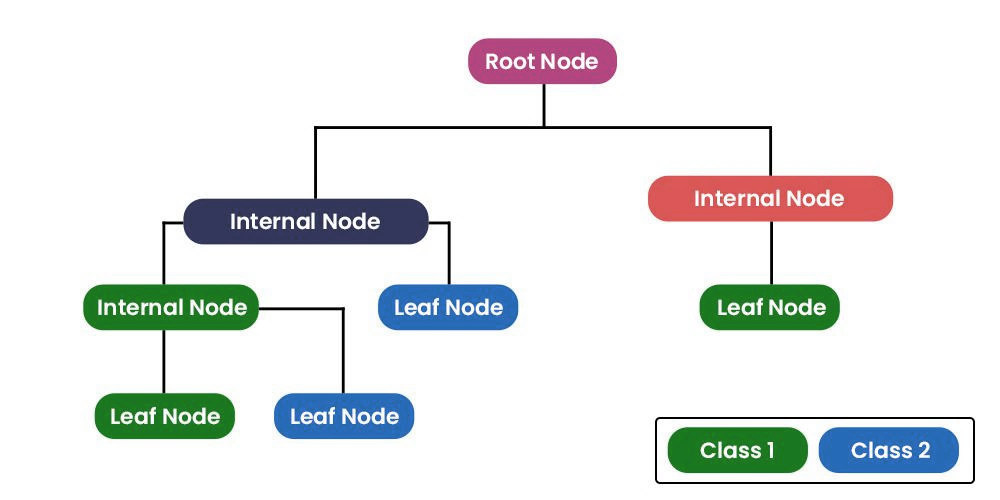


Figure 13. Decision Tree algorithm.(13)

#### K-Nearest Neighbor (K-NN)

K-nearest neighbor (K-NN) stands out as an intuitive supervised learning algorithm, and relatively straightforward classification algorithm that relies on proximity metrics. However, its efficacy is heavily influenced by factors such as the choice of the parameter k, the method used for distance calculation, and the selection of appropriate predictors.

K-NN's capacity to classify unlabeled observations by placing them in the same class as the most similar labeled samples is one of its primary features. However, because it stores every training instance, its simplicity comes at a cost: when working with large training sets, it becomes computationally expensive in terms of both time and space. K-NN is a nonlinear classifier that can precisely assess decision boundaries in spite of its simplicity.

In a study by Wang et al., four emotions (joy, sadness, pleasure, and anger) were classified with a mean recognition rate of 82% utilizing a combination of K-NN with feature selection using a Tabu search heuristic method and 4-fold cross-validation (1). Using a variety of algorithms, including as K-NN, neural networks, LDA, and QDA, in addition to feature selection approaches like SFS and SBS and incorporating multiple cross-validation strategies, Kolodyazhniy conducted pattern classification analysis. An accuracy rate of 77.5% for unknown people in unfamiliar settings was attained by the K-NN model, specifically with a 17-fold cross-validation procedure, which showed superior accuracy in both participant- and stimuli-dependent categorization(1).(Figure 14)

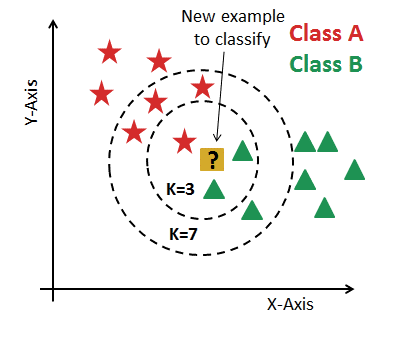


Figure 14. K Nearest Neighbor algorithm.(14)

In summary, while K-NN offers simplicity and interpretability, careful consideration of parameters and feature selection techniques is crucial for optimizing its performance in emotion recognition tasks. Additionally, its ability to accurately classify data points based on their proximity to labeled examples makes it a valuable tool in pattern recognition tasks.

### Deep Learning Methods and Results

#### Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) represent a specialized type of deep learning model designed for supervised learning tasks, most likely into fields like image classification and object detection. Within CNN, convolutional layers are like filters that slide over the input data to detect patterns. Each filter learns to recognize a specific feature, like edges, corners, or textures, and as the network goes deeper, these filters combine to detect more complex features, like shapes or objects.

Pooling layers are so important in reducing the dimensionality of the feature maps while preserving important information. They do this by summarizing groups of pixels, effectively down sampling the data. Fully Connected Layers take the features extracted by convolution and pooling and use them to make predictions(15). They connect every neuron from the previous layer to every neuron in the next layer, allowing the network to learn complex relationships between features. In the classification part of the CNN, we have several components that come into play:

* Input Layer: This is the initial layer where the raw data, like an image, is fed into the network.
* Multiple Hidden Layers: These layers consist of convolutional and pooling layers that extract and refine features from the input data.
* Output Layer: The network's last layer, which is where the predictions are made. In this layer, every neuron corresponds to a distinct class, and the anticipated class is indicated by the neuron with the highest activity value.

CNNs exhibit effectiveness because they learn to recognize features automatically from the data, thanks to the sharing of weights and local connections(1). This mechanism helps to reduce the complexity of the network while maintaining accuracy. (Figure 15)

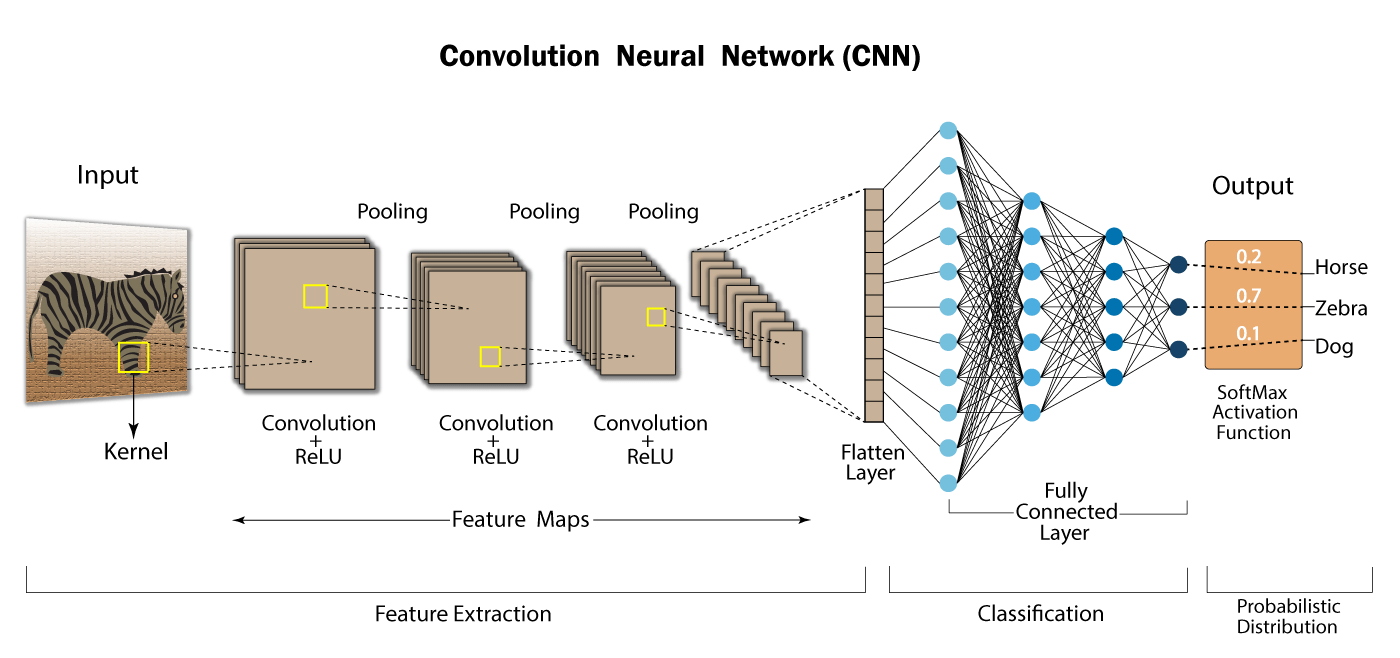


Figure 15. Convolutional Neural Network layout.(16)

#### Deep Belief Networks (DBN)

Often called a generative graphical model, deep belief networks (DBNs) are a kind of deep neural network that function as a probabilistic generative model. DBNs are not like standard feedforward neural networks; instead, they are made up of several layers of hidden units, or latent variables, connected only between layers but not within them. Pattern recognition in the input data is done by lower-level restricted Boltzmann machines (RBMs) in a DBN architecture. The output of these RBMs is then used as input by higher-level RBMs. The input data is used by each layer of the DBN to learn progressively more intricate features.

Zheng et al. have presented a recent improvement in DBNs that uses EEG data to distinguish between positive and negative emotions by adding characteristics based on differential entropy (DE). Zheng conducted studies wherein the DBN classifier outperformed other classifiers, including Support Vector Machine (SVM) and K-Nearest Neighbors (K-NN), in correctly classifying three distinct emotional states: positive, negative, and neutral. (Figure 16)(1).

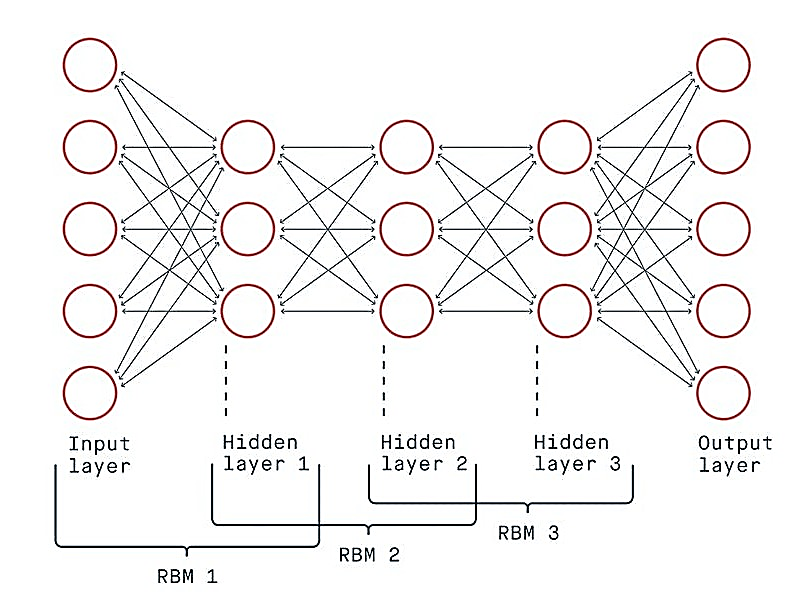


Figure 16. Deep Belief Network feed-forward architecture layout.(17)

#### Deep Neural Networks (DNN)

Each layer of a deep neural network (DNN) is made up of several neurons that work together to create a hierarchical structure. Each layer in the cascade learns more complicated patterns in the incoming data since the output of the previous layer is the input of the next layer, and so on. Deeper layers often learn high-level abstractions in the data, while lower levels typically learn low-level features. Because DNNs function as feed-forward neural networks, they have the simplest structure of any neural network type. This indicates that information flows from the input layer to the output layer via the hidden layers in a single direction. (Figure 17)(15).

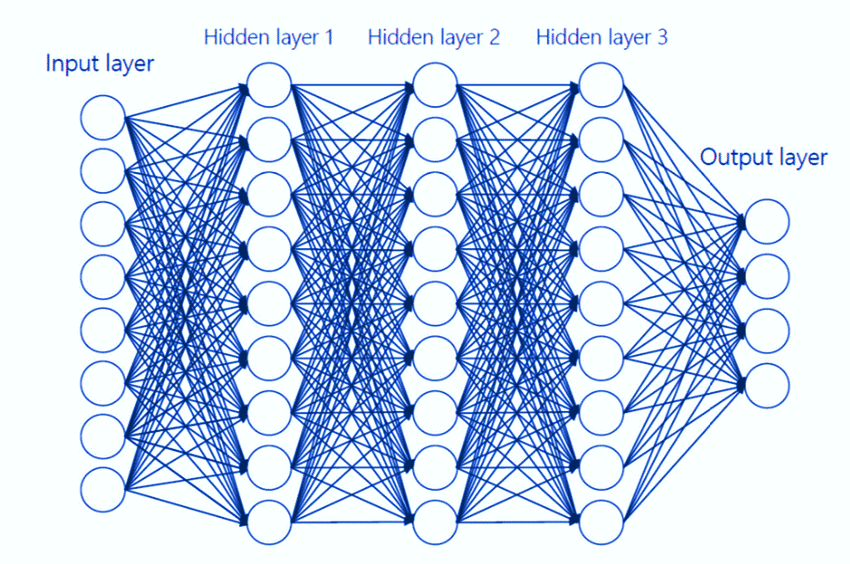


Figure 17. Deep Neural Network feed-forward architecture layout.(18)

#### Probabilistic Neural Networks (PNN)

A deep, feed-forward neural network is the Probabilistic Neural Network (PNN). Because of its simple structure and quick learning speed, it can achieve high accuracy and improved resistance to mistakes and noise. An input layer, a pattern layer, a summation layer, and a decision layer are the four layers that make up a PNN. Most terms are computed by the PNN using the training set of terms (19).

While the pattern layer recognizes and records the underlying patterns in the data, the input layer handles the incoming data. These patterns are combined in the summation layer, and then probabilistic outputs are generated in the output layer, enabling a more sophisticated comprehension of the predictions. (Figure 18) Because of its unique structure, PNN performs exceptionally well in tasks that involve uncertainty, which makes it especially useful in real-world applications.Top of Form

In a study conducted by Zhang et al., PNN was employed for emotion recognition using EEG data, with a focus on extracting features related to sub-band power. Results indicated that PNN achieved slightly lower classification rates compared to SVM. Specifically, for arousal classification, PNN attained 81.76% accuracy, while SVM reached 82%. Similarly, for valence classification, PNN achieved 81.21% accuracy, whereas SVM attained 82.26%.(1)

However, it's worth noting that PNN required fewer channels to achieve comparable performance to SVM. For arousal classification, PNN utilized only 9 channels compared to SVM's 14, and for valence classification, PNN used 9 channels while SVM required 19. This suggests that PNN may offer more efficient resource utilization while still maintaining competitive performance in emotion recognition tasks based on EEG data.

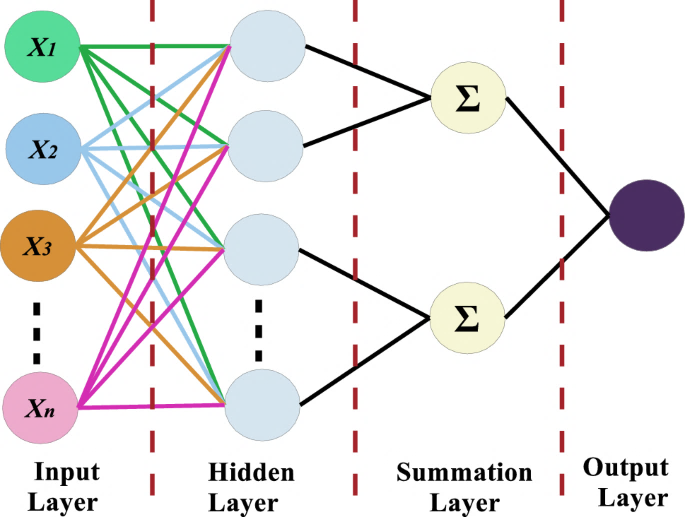


Figure 18. Probabilistic Neural Network architecture layout.(20)

## Datasets

### DEAP *---* Database for emotion analysis using physiological signals

The DEAP dataset, developed by Koelstra et al., stands as a pivotal resource in emotion analysis research. It comprises diverse physiological signals, including 32-channel EEG data and 12 additional signals such as EMG, RSP, ERG and GSR. These signals were collected while 32 participants viewed 40 music videos and rated them based on arousal, valence, dominance, like/dislike, and familiarity.(21) This rich dataset facilitates the exploration of emotions through various modalities, providing researchers with valuable insights into human emotional responses. The EEG data was sampled at 512 Hz and pre-processed by down sampling to 128 Hz, with EOG artifacts removed using a band-pass filter.

Researchers have leveraged the DEAP dataset to achieve remarkable accuracy in emotion recognition tasks. Using advanced classification algorithms and feature sets, they have accurately identified emotions such as anger, joy, and surprise.(1) Additionally, the dataset offers subjective ratings from participants, enabling a deeper understanding of emotional experiences. With its comprehensive collection of physiological signals and subjective ratings, the DEAP dataset serves as a cornerstone for studying emotions and advancing emotion analysis techniques.

Using the DEAP dataset, Bălan et al. studied emotion recognition and obtained impressive accuracy rates on a variety of emotions (Figure 19). Using SVM with Petrosian and Higuchi fractal dimensions features, the highest accuracy rates were, for example, 98.02% for anger, 95% for disgust, and 100% for joy by LDA with EEG raw data figures; 96% for surprise, SVM and SFS with EEG raw values; 90.75% for fear, k-NN with EEG raw figures without feature selection; and 90.08% for sadness, SVM with Higuchi fractal dimensions without feature selection (1). Various classification algorithms and feature sets were also used to accurately identify other emotions, including surprise, fear, and melancholy.

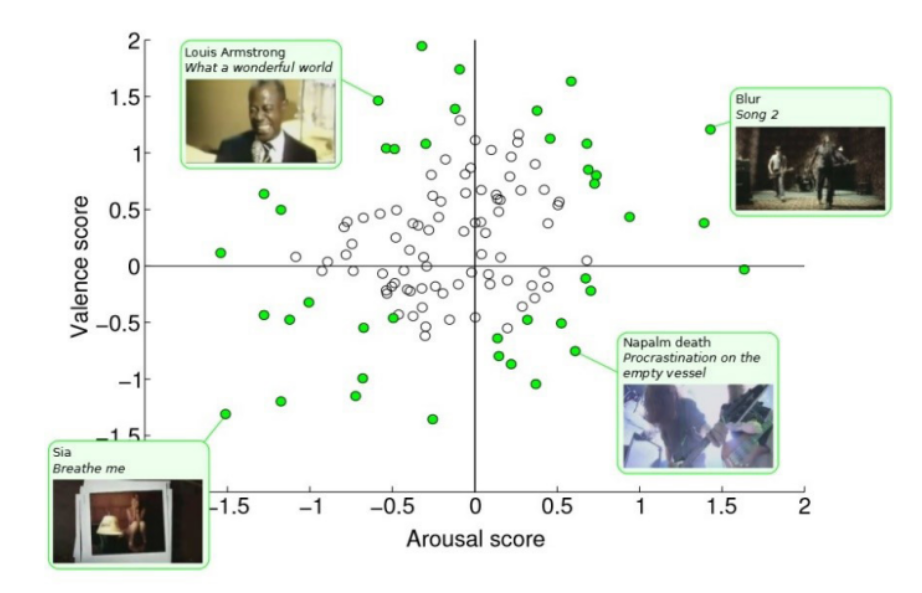


Figure 19. Video clip valence and arousal values in the DEAP database. Videos that are included are displayed in green.(21)

### SEED *---* SJTU emotion EEG dataset

The SEED database was presented by Zheng and Lu. It is made up of EEG and facial video recordings of 15 subjects showing three different emotions: happy, neutral, and negative (21). The dataset consists of 45 experiments, with 15 trials consisting of 4-minute movie snippets in each experiment. EEG signals were band-pass filtered in the frequency range of 0–75 Hz after being collected from 62 channels via a 1000 Hz sampling frequency and then down sampled to 200 Hz. Notably, the SEED database is well known for its efficacy at identifying emotions from EEG signals. Differential Entropy (DE) features are widely used, especially in the 5-frequency bands.

Differential Entropy-Pearson Correlation Coefficient Matrix (DE-PCCM), a unique feature extraction technique, in conjunction with CNN produced a maximum classification accuracy of 98.59% in one study by Li et al (1). Another study computed spectral Shannon and K-nearest neighbor (K-NN) entropies using an empirical wavelet transform based on the Fourier-Bessel series expansion, yielding a classification accuracy of 94.4%.(1)

Moreover, emotional EEG signals and eye movement signals for four and five distinct emotions, respectively, are contained in SEED-IV and SEED-V, the next evolution of the SEED database. Happy, sad, fear, and neutral emotions are all included in SEED-IV, which uses 72 movie clips as stimuli. In contrast, SEED-V uses 15 film clips as stimuli and covers happy, sad, fear, disgust, and neutral feelings. Participants in both databases complete several sessions while exhibiting a range of emotions. Band-pass filtering between 1 and 75 Hz is used in preprocessing to get rid of artifacts and noise.

Using bimodal deep auto-encoder approaches and deep canonical correlation analysis on SEED, SEED-IV, and SEED-V, notable recognition accuracies were obtained. Furthermore, in order to gauge vigilance levels, SEED-VIG, a variant of SEED, captures EEG and EOG signals during a simulated driving scenario. (1) These databases serve as invaluable resources for studying emotions and developing emotion recognition systems, providing access to raw and preprocessed signals for research purposes. More information can be found on the respective database websites.

### DREAMER

A multimodal database called DREAMER was created to investigate emotions using EEG waves. It consists of 14 channels of recorded EEG data from 23 patients, 14 of whom were male and 9 of whom were female. Eighteen different emotions are targeted by the audio-visual stimuli in the form of film clips in the database: amusement, anger, tranquility, disgust, excitement, fear, happiness, sadness, and surprise. Each film clip lasts anything from 65 to 393 seconds.(1)

Using a low-cost, portable, wearable, wireless technology, such as the Emotiv EPOC system and SHIMMER wireless sensor, all EEG signals in the DREAMER database were recorded. According to the International 10–20 system, the electrodes, which included AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, M1, and M2, were arranged. Signals were recorded at a sampling rate of 128 Hz, providing researchers with rich data to explore various aspects of human emotion. For detailed information on the protocol and usage of the database, researchers can refer to the provided resources.(1)

### MAHNOB HCI- tagging database

The MAHNOB-HCI database was created by Soleymani et al. in 2012. It is a multimodal dataset that includes measurements of skin temperature, respiration amplitude, 32-channel EEG signals, 3-channel ECG signals, 2-channel ERG, and 2-channel GSR. Twenty-seven participants provided data while seeing pictures and videos. (21).

The study was divided into two sessions. In the first, participants watched video clips and then answered an emotional questionnaire right away. They saw brief films and images in the second session, both labeled and unlabeled, and their emotional responses were recorded. Every signal was captured using the 10-20 international standard system.

The International Affective Picture System (IAPS) offers natural emotional responses to photos in the MAHNOB-HCI database, an open-source resource. It includes facial expressions from the 27 participants, audio recordings, ocular gaze data, and responses from 32-channel EEG signals (1). The database includes results from two experiments: participants in the first saw 20 films and images that evoked emotions, and in the second, they saw 28 photos and 14 brief video clips that included both tagged and untagged human gestures. On a scale of 1 to 9, participants assessed the stimuli for valence, arousal, dominance, and agreement with the displayed tags.

Koelstra and I. Patras combined regression-estimated weights fusion (W-REG) with recursive feature elimination (RFE) to obtain valence and arousal classification accuracies of 73% and 72.5%, respectively. The Probabilistic Neural Network (PNN) achieved the best accuracy of 96.9% by carefully choosing the right features and channels using evolutionary computation algorithms.(1)

## Metrics

The measures used for the for the calculation of the performance metrics (22) are:

* **tp** – true positive,
* **fp** – false positive,
* **fn** – false negative,
* **tn** – true negative counts.

The different performance metrics used are accuracy, precision, true positive rate, false positive rate, f1 measure, kappa score, area under curve.

**Accuracy:** The percentage of correctly identified inputs relative to the total number of inputs is known as accuracy. It is computed mathematically as the total of true positives, or occurrences that were correctly classified, divided by the total of true positives plus the total of false negatives, or instances that were wrongly classified. As a result, the number of accurate labels for various classes is not distinguished by accuracy, the most widely used empirical metric.

|  |  |  |
| --- | --- | --- |
|  |  | Eq.1 |

**Precision:** The precision of each class's positive predictions is measured. It is the ratio of true positives to the total of false positives (incorrectly categorized as positive) for a certain class. The average precision value in multi-class classification is calculated by summing the precision values for all the classes.

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 2 |

**True Positive Rate/Recall:** Recall, another name for True Positive Rate, is a measurement of the proportion of properly identified occurrences of a given class relative to all instances of that class. The ratio of true positives to the total of false negatives and true positives is used to compute it.

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 3 |

**False Positive Rate/Specificity:** The percentage of occurrences of a particular class that are wrongly classified relative to all instances, including instances of other classes, is known as the False Positive Rate. The average false positive rate in multi-class classification is calculated by summing the false positive rates for all classes.

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 4 |

**F1 Measure:** The F1 Measure is a metric used to evaluate accuracy that takes precision and recall into account. It is determined by taking the harmonic mean of recall and precision, yielding a single score between 0 and 1. A score of 0 denotes an inaccurate categorization, whereas a score of 1 denotes perfect classification.

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 5 |

**Kappa Score:** Kappa Score evaluates the agreement between observed and expected classes, accounting for chance agreement. It is computed using the formula (accuracy - expected accuracy) divided by (1 - expected accuracy), where expected accuracy is calculated based on the observed and expected classes.

**Area-under-curve (AUC):** The receiver operating characteristic (ROC) curve is used to calculate the area under the curve (AUC) for a machine learning classifier. Higher values indicate better classification ability. AUC stands for the area under the ROC curve. The AUC of a perfect classifier would be 1, but the AUC of a random classification would be 0.5.

**Confusion Matrix:** A confusion matrix summarizes a classification model's performance is called a confusion matrix. There are four categories in it. Beyond simple measurements like accuracy or precision, the confusion matrix offers a holistic picture of the model's performance and enables a deeper study. It's especially helpful for pinpointing the precise kinds of mistakes the model makes and comprehending its advantages and disadvantages for various class settings. (Table 2)

Table 2. Confusion matrix sample.

|  |  |  |
| --- | --- | --- |
| *Class/ Recognized* | *As Positive* | *As Negative* |
| Positive | tp | fn |
| Negative | Fp | tn |

# CHAPTER 3

# METHODOLOGY

## Datasets Used in This Research

### Dreamer Dataset

Electroencephalogram (EEG) and electrocardiogram (ECG) signals acquired during affect elicitation using audio-visual stimuli are stored in a multi-modal database called DREAMER(23). The recordings from 23 participants are included in the data, along with their ratings of the valence, arousal, and dominance of their own affective states after each stimulation. Portable, wearable, wireless, inexpensive, off-the-shelf technology that is appropriate for incorporating affective computing into daily applications was used to acquire the signals. EEG and ECG data were specifically collected using the Emotiv EPOC wireless EEG headset and the Shimmer2 ECG sensor, respectively. (23)

The DREAMER database includes participant ratings and physiological recordings from a study in which eighteen film segments selected and assessed by Gabert-Quillen et al. were seen by twenty-three volunteers. EEG and ECG data were recorded during the experiment, and participants used five-point rating scales to indicate how arousing, valence, and dominant they felt their emotional responses to be.

The DREAMER database is stored in a file named "DREAMER.mat" in Matlab format.(23) Upon loading this file, a variable named "DREAMER" is instantiated in the workspace. The structure of the "DREAMER" variable is as follows:Top of Form

DREAMER = struct with fields:

* Data: {1×23 cell}
* EEG\_SamplingRate: 128
* ECG\_SamplingRate: 256
* EEG\_Electrodes: {'AF3' 'F7' 'F3' 'FC5' 'T7' 'P7' 'O1' 'O2' 'P8' 'T8' 'FC6' 'F4' 'F8' 'AF4'}
* noOfSubjects: 23
* noOfVideoSequences: 18
* Disclaimer: 'While every care has been taken…'
* Provider: 'University of the West of Scotland'
* Version: '1.0.2'
* Acknowledgement: 'The authors would like to thank…' (23)

The cell DREAMER.Data{i} contains the data for the ith participant and is structured as follows: struct with fields:

* Age: 'X'
* Gender: 'X' ('male' or 'female')
* EEG: [1×1 struct]
* ECG: [1×1 struct]
* ScoreValence: [18×1 double]
* ScoreArousal: [18×1 double]
* ScoreDominance: [18×1 double]

The EEG and ECG recordings are stored in the DREAMER.Data{i}. EEG and DREAMER.Data{i}. ECG variables respectively which are structured as follows:

baseline: {18×1 cell}

stimuli: {18×1 cell}

The "baseline" variable in the "DREAMER" structure contains the recordings for the neutral clip that is played prior to each film clip, while the recordings for the stimuli film clips are kept in the "stimuli" variable. In particular:

* Each cell "baseline{i}" contains the data for the neutral clip shown before the ith film clip.
* Each cell "stimuli{i}" contains the data for the ith film clip stimulus.

Table 3. jth column of EEG recordings corresponds to the following EEG positions.

|  |  |
| --- | --- |
| j | Position |
| 1 | AF3 |
| 2 | F7 |
| 3 | F3 |
| 4 | FC5 |
| 5 | T7 |
| 6 | P7 |
| 7 | O1 |
| 8 | O2 |
| 9 | P8 |
| 10 | T8 |
| 11 | FC6 |
| 12 | F4 |
| 13 | F8 |
| 14 | AF4 |

For EEG data:

* Each EEG recording is represented as an M x 14 matrix, where M is the number of available samples.
* Each column of the matrix contains samples from one of the 14 EEG channels.

The 14 EEG channels correspond to specific electrode positions. (23) The jth column of the EEG recordings corresponds to the following electrode positions (Table 3).

## Data Preprocessing

Preprocessing is one of the crucial steps one should do on their data. It essentially is the preparation and transformation of the data in a more suitable way for feature extraction and further for data mining(24). While doing data preprocessing is worth mentioning that we do not affect the original data, rather we take it and work with it in a way that it will best fill our needs. There are several steps and techniques you can do to preprocess the data but the ones we have chosen are: Bandpass filtering, Power Spectral Density Calculation, Baseline- Stimulus Ratio Calculation and Feature Normalization.

### Bandpass Filtering (BPF)

Bandpass filtering is an important step in EEG signal processing. Specific frequency bands have been isolated by us for further analysis. The overall goal of the bandpass filter is to isolate the important frequency components by eliminating high-frequency noise and low-frequency drift. Next, for analysis, we have selected the following bands:

* **Delta band:** Prominent during deep sleep
* **Theta band:** Associated with drowsiness, relaxation and light sleep.
* **Alpha band:** Linked with relaxation and mental coordination, helping to understand brain states during rest and calm.
* **Beta band:** Associated with active thinking and focus.
* **Gamma band:** Linked to high-level functioning and information processing.

### Power Spectral Density Calculation (PSD)

We have then calculated the Power Spectral Density (PSD) for each filtered signal. PSD measures the power of each frequency component within the signal, which is essential for understanding the distribution of power across different frequency bands. This is done using the Welch method, which is a robust technique for estimating the power of a signal.

### Baseline-Stimulus Ratio Calculation

The Baseline-Stimulus Ratio Calculation is indeed a crucial step in analyzing EEG data. It helps to address individual differences and refine the results by quantifying the change in EEG activity induced by a stimulus. This calculation involves comparing the Power Spectral Density (PSD) of EEG signals during baseline (resting) conditions with the PSD during stimulus (task) conditions. By computing the ratio of the mean PSD during the stimulus condition to the mean PSD during the baseline condition for each frequency band, we can specifically highlight the effects of the stimulus on brain activity relative to the baseline state. This normalization process enhances the interpretability of the results, providing valuable insights into the neural responses to different stimuli.

### Scaling

A crucial preprocessing procedure used in deep learning techniques to standardize the range of features or independent variables in the data is scaling. Its importance is in making sure that, because of scale disparities, no specific characteristic has an undue impact on the model's training process. The provided code fits the scaler to the training set and then transforms it using the fit\_transform() method of the scaler on the training data (X\_train). Each feature in the training data is calculated for mean and standard deviation in this procedure, and the results are then utilized to scale the data properly. After that, the scaling parameters discovered from the training data are used to apply the transform () technique to the test data (X\_test).

Maintaining consistency by using the same scaler instance for both training and test data is crucial to ensure uniformity in scaling across the dataset. This standardized data preprocessing step optimizes the performance of deep learning models by facilitating more effective learning and generalization.

### Feature Normalization

After extracting the features, we normalized them using standard scaling to make sure every feature adds the same amount to the analysis. The values are normalized to have a zero mean and a one standard deviation. Ensuring that every feature has an equal impact on the model is a crucial step for machine learning algorithms that are sensitive to the volume of input data.

## Feature Extraction

In order to make sure that our data is less complex, feature extraction is a critical phase that lowers the cost of our data mining process and the risk of losing valuable information. We can provide the most accurate description of a large amount of data by using feature extraction. We have a few different kinds of features because it's not always straightforward to locate the characteristics that are most suited. EEG analysis uses three types of features: time domain, frequency domain, and time-frequency domain (4).

### Frequency Domain

Analyzing the signals of data in terms of frequency is what we call a “frequency domain” (4). In order to sample a signal without aliasing, we first applied bandpass filters to extract out specific EEG frequency bands, which are: delta, which has a frequency of 0.5–4Hz; theta, which has a frequency of 4–8Hz; alpha, which is 8–13Hz; beta, which is 13–30Hz; and gamma, which is up to 64Hz because our data sampling rate is 128Hz. This is because, according to the Nyquist theorem, the sampling frequency must be at least twice that of the highest frequency on the signal. Using the Weltch approach, we have calculated the Power Spectral Density (PSD) following bandpass-filtering. This method determines the signal's strength across several frequency ranges. Finally, we have taken the mean of the PSD values for each frequency band and used them as features for the frequency domain.

### Time Domain and Statistical

Analyzing the signals of data in terms of time is what we call a “time domain” analysis. We have focused on extracting statistical features from the time domain signals of EEG data. These features summarize the distribution and shape of the signals over time, providing insights into the underlying characteristics of the EEG data. By analyzing these statistical features, we can capture important temporal patterns and variations in the EEG signals. These features were computed for both the baseline and stimulus periods, and the stimulus features were normalized by the baseline features to highlight changes due to stimuli. This approach allows us to effectively utilize time domain analysis to understand the dynamics of EEG signals. There are different examples of time domain statistical features but what we have used are the mean, kurtosis, skewness, standard deviation.

#### Mean

The total of all the values in a dataset divided by the total number of values is the mean, or average. It offers a measurement of the data's central tendency. In this instance, the mean of the EEG signals can represent the general level of brain activity; a greater mean denotes more intense activity. The equation (where x\_k denotes the k-th observation and N is the total number of observations):

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 6 |

#### Standard Deviation

Standard deviation measures the amount of variation in a dataset. For EEG signals, the standard deviation can sow how complex are the brain functions or responses where the higher the variability the more complex they are. The formula (N is the number of observations and is the k-th observation):

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 7 |

#### Skewness

The skewness of a dataset indicates how asymmetrically its values are distributed. The longer tail on the right side of the distribution is shown by a positive skewness, whereas the longer tail on the left is indicated by a negative value. According to our research, skewness can show whether brain activity is high or low. For instance, a positive skewness may suggest that there are sporadic spikes in brain activity. It aids in our comprehension of whether the brain experiences high or low activity states more frequently. The equation (where denotes the k-th observation and N is the total number of observations):

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 8 |

#### Kurtosis

Our understanding of the frequency of outliers in the distribution of our values inside a dataset is aided by the Kurtosis. When the kurtosis of our data is high, it indicates that there are many outliers in it, and when it is low, there are few outliers. In our situation, a high Kurtosis may potentially indicate the existence of uncommon but important events. The equation (where denotes the k-th observation and N is the total number of observations):

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 9 |

## Machine Learning Methods used in this Research

Valence, arousal, and dominance levels from EEG data have been classified using machine learning techniques. Finding patterns and creating data-driven forecasts depend heavily on machine learning. We give a summary of the techniques we have employed below.

### Support Vector Machine (SVM)

As we've already discussed, the way this strategy operates is by identifying a hyperplane in our feature space that efficiently divides data points into different classes. In this work, a grind search cross-validation was used to apply the SVM model. We were able to test many combinations of hyperparameters in this procedure, including our kernel type (linear, polynomial, or radial basis function), the kernel coefficient (gamma), and the regularization parameter (C). The accuracy, precision, recall, and F1-score of the trained SVM model were assessed using datasets from both training and testing. These metrics aid in our comprehension of the model's functionality.

### Random Forest (RF)

Through the use of many decision trees, the Random Forest model is a learning technique that can handle tasks involving both regression and classification. In our case, the Random Forest was established, and the EEG dataset was used to train it. Additionally, grind search cross-validation was employed to ascertain the optimal parameters for the RF. We experimented with a variety of factors, including the total number of trees in the forest, the tree's maximum depth, the minimum number of samples needed at each leaf node, and the minimum number of samples needed to split a node. The accuracy, precision, recall, and F1-score of the trained RF model were assessed using datasets from both training and testing.

### Decision Tree (DT)

With each node representing a feature, branches representing decisions, and leaf nodes representing outcomes, a decision tree can classify data by dividing it depending on attribute values. The EEG dataset has been used for initialization and training. The best hyperparameters for the Decision Tree model were found using the grind search cross-validation in this model. We experimented with various values for the maximum depth of the tree, the minimum samples needed to split a node, and the minimum samples needed at a leaf node to determine which combination of these parameters produced the highest accuracy. The accuracy, precision, recall, and F1-score of the trained DT model were assessed using datasets from both training and testing.

### K-Nearest Neighbor (K-NN)

Data points are categorized in K-NN according to how close they are to labeled instances. The selection of k determines how well it performs. We performed parameter optimization, examining several values of k to determine the ideal number of neighbors, in order to determine the best choice for k. Next, we train the KNN model for this value using this k. After that, analyze the training and testing datasets to determine the F1-score, accuracy, precision, and recall.

## Deep Learning Methods used in this Study

Using EEG data, we have applied deep learning techniques to classify dominance, arousal, and valence levels. Deep learning is a branch of machine learning that specializes in extracting complicated patterns and representations from large amounts of data. We provide a summary of the deep learning techniques we used in our investigation below.

### Conventual Neural Networks (CNN)

Convolutional neural networks, or CNNs, are typically used for image recognition applications, but they can also be used for text processing and other sequential data analysis tasks, such as time series analysis. A 1D CNN architecture was used in this work, taking use of its versatility outside of traditional image-based applications. Convolutional layers were the first layers in the CNN model architecture, and then fully connected layers. The next set of fully connected layers made classification based on the learnt features easier, and these convolutional layers were crucial in removing hierarchical features from the input sequences. Specifically, different kernel sizes and numbers of filters were tested in the convolutional layers in an effort to capture the various degrees of abstraction present in the input data.

### Recurrent Neural Network

Due to their ability to process sequential data by maintaining an internal state or memory, recurrent neural networks (RNNs) are especially well-suited for tasks involving temporal relationships, like sentiment analysis and time series prediction. This work built a specific RNN architecture using Long Short-Term Memory (LSTM) units. Beyond the limits of conventional RNNs, LSTMs are well known for their capacity to capture long-range dependencies.

The RNN model successfully identified patterns in sequential data by utilizing LSTM units that were furnished with memory cells that could hold information over time steps. In order to optimize the RNN model's performance for the particular binary classification problem under research, the investigation involved exploring a variety of configurations, including varying numbers of LSTM layers, dropout rates, and activation functions.

### Deep Neural Network

Deep Neural Networks (DNNs), also referred to as feedforward neural networks, constitute fundamental architectures in the realm of deep learning. Comprising multiple layers of neurons interconnected in a feedforward manner, DNNs are adept at tackling a wide array of tasks, including binary classification. The DNN model architecture employed in this study consisted of densely connected layers, complemented by dropout regularization to mitigate overfitting risks.

Dropout, a regularization technique that randomly nullifies input units during training, was instrumental in promoting model generalization by preventing excessive reliance on specific features. Throughout the investigation, diverse configurations of hidden layers and activation functions were explored to identify the optimal architecture tailored to the binary classification task.

## Metrics used in this Study

Metrics are a means to evaluate the combined effectiveness of data learning and machine learning models in identifying dominance, valence, and arousal from EEG data. They aid in our understanding of the models' capacity to accurately categorize situations in various emotional states. I have employed the following metrics: ***accuracy, precision, recall,*** and ***f1-score***. To facilitate our comparison, we have combined all of the model findings into a single Excel table.

# CHAPTER 4

# RESULTS AND DISCUSSION

In this section, we further advance our understanding of EEG-based emotion recognition, employing a diverse array of analytical methods to unveil the intricate workings of human emotional responses. Our investigation begins by examining the frequency-domain data, where we undertake both binary (0 and 1) and ternary (0, 1 and 2) classification tasks associated with valence, arousal, and dominance. Next, we have shifted our focus to the time-domain, isolating important features to gain deeper insights into the dynamic nature of emotional processing.

In our pursuit of understanding, we embrace a holistic approach, employing both conventional machine learning techniques such as Random Forest, Decision Trees, Support Vector Machines (SVM), and k-Nearest Neighbors (k-NN), as well as advanced deep learning methodologies like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Deep Neural Networks (DNNs). This combination of traditional and cutting-edge methods enriches our analysis, offering nuanced perspectives on feature extraction and modeling strategies, and ultimately advancing our comprehension of how EEG signals encode emotional states.

## Frequency Domain Results on Machine Learning Algorithms

To separate particular EEG frequency bands (delta, theta, alpha, beta, and gamma) for the frequency-domain analysis, we used bandpass filtering. For each band, we calculated the Power Spectral Density (PSD). To ensure the reliability of our models, we applied normalization to the PSD features, standardizing them to have zero mean and unit variance. This step was crucial in ensuring that each feature contributed equally to the model's performance, avoiding biases due to different feature scales.

We first used a binary technique to classify valence, arousal, and dominance, splitting each dimension into two levels. For instance, emotional valence was divided into low and high states, which corresponded to negative and positive emotional valence. In a similar manner, binary classification was used to identify low and high levels of dominance and arousal, respectively. We expanded our research to include ternary classification in addition to binary classification, which provided a more sophisticated understanding of the emotional states stored in EEG signals.

In order to ensure that all classes were fairly represented in the training data, we used random oversampling to correct class imbalances that were present in the binary and ternary classification scheme. In order to ensure equity in class representation, this strategy rectified the imbalance by replicating instances from the minority class. This oversampling strategy enhanced the robustness of our models, allowing them to generalize more effectively across all classes.

To categorize based on the extracted and normalized PSD characteristics, we used a number of machine learning methods, such as Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), and K-Nearest Neighbor (K-NN). The models were assessed and trained using common metrics like F1-score, recall, accuracy, and precision. Below we have the complete results of binary and ternary classification. To further enhance model performance, hyperparameters for each model were optimized using grid search cross-validation.

### Binary Classification

* Valance results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 4. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 100 |
| max\_depth: 5,10,20 | 20 |
| min\_samples\_split: 2,5,10 | 10 |
| min\_samples\_leaf: 1,2,4 | 1 |

Table 5. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 6.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 5 |
| min\_samples\_leaf: 1,2,4 | 2 |



Figure . KNN performance for different values of k where the best one is k=3.

Table 7. Table with results of valance for frequency domain, binary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 99.71751 | 73.68421 | 73.80952 | 73.56202 | 73.56981 | 0.235219 | 0.033426 |
| SVM | 81.63842 | 73.68421 | 73.80952 | 73.56202 | 73.56981 | 0.012033 | 0.022055 |
| DT | 95.19774 | 80.92105 | 80.91729 | 80.89051 | 80.90039 | 0.011539 | 0.009423 |
| KNN | 89.54802 | 83.55263 | 83.96193 | 83.69716 | 83.53482 | 0.003638 | 0.012096 |

* Arousal results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 8. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 300 |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 2 |

Table 9. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 10.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 1 |

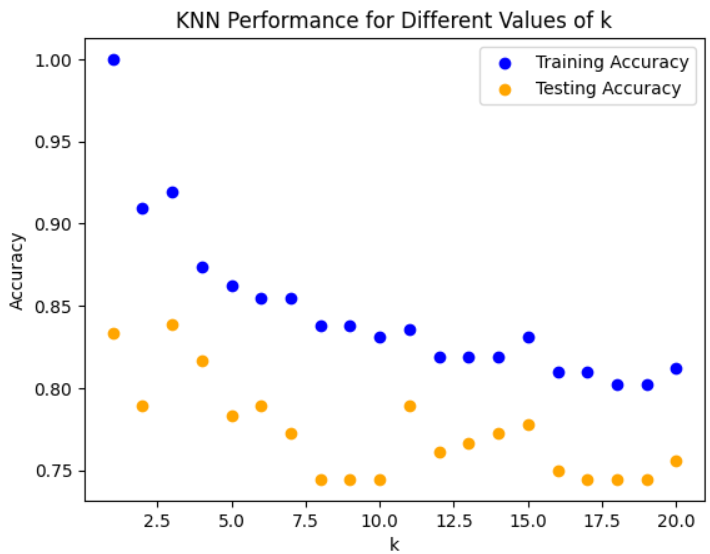


Figure 21. KNN performance for different values of k where the best one is k=3.

Table 11. Table with results of arousal for frequency domain, binary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 98.4 | 93.88889 | 94.04762 | 93.85727 | 93.87963 | 0.770879 | 0.080607 |
| SVM | 85.95238 | 82.22222 | 82.64194 | 82.15829 | 82.14286 | 0.010918 | 0.021215 |
| DT | 98.57143 | 87.22222 | 89.44328 | 87.091 | 87.01013 | 0.013937 | 0.01084 |
| KNN | 91.90476 | 83.88889 | 86.32376 | 83.74491 | 83.57199 | 0.004278 | 0.015696 |

* Dominance results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 12. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 100 |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 5 |
| min\_samples\_leaf: 1,2,4 | 1 |

Table 13. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 14.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 20 |
| min\_samples\_split: 2,5,10 | 5 |
| min\_samples\_leaf: 1,2,4 | 1 |

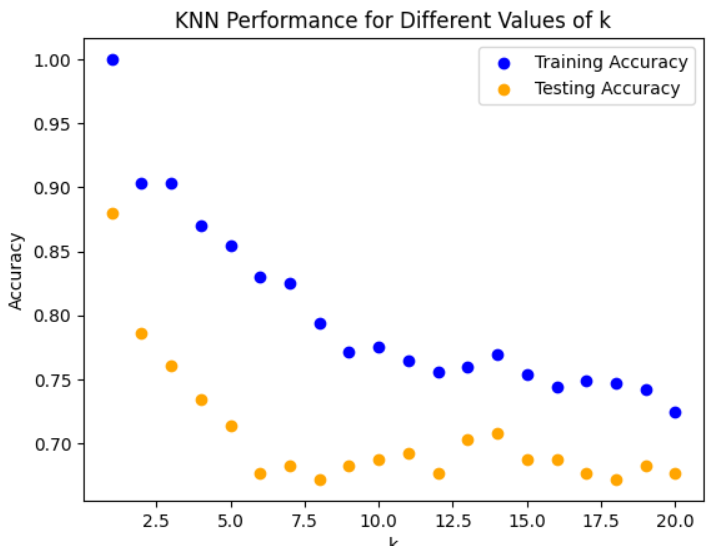


Figure 22. KNN performance for different values of k where the best one is k=2.

Table 15. Table with results of dominance for frequency domain, binary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 98.5 | 89.58333 | 89.94755 | 89.68736 | 89.57315 | 0.266196 | 0.038658 |
| SVM | 95.06726 | 83.33333 | 83.80869 | 83.45636 | 83.30435 | 0.012171 | 0.02616 |
| DT | 99.10314 | 85.9375 | 86.75251 | 86.09422 | 85.89119 | 0.014785 | 0.010731 |
| KNN | 90.35874 | 78.64583 | 82.89037 | 79.0165 | 78.07426 | 0.005196 | 0.017164 |

### Ternary Classification

* Valance results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 16. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 100 |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 2 |

Table 17. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 18.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 20 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 2 |

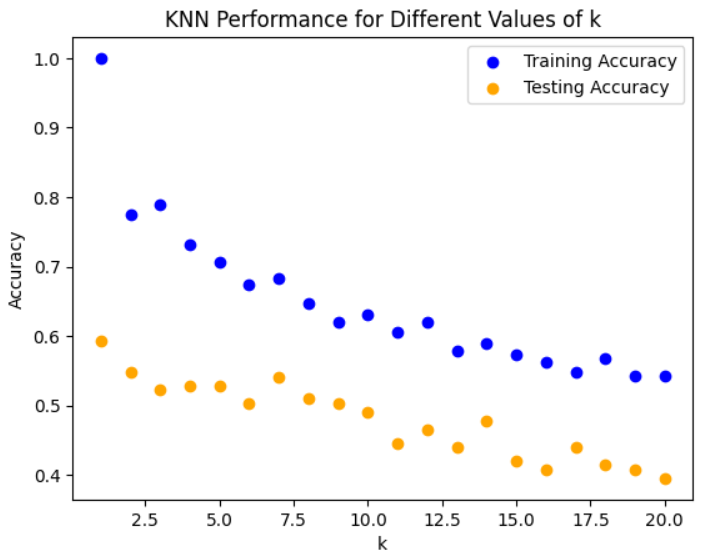


Figure 23. KNN performance for different values of k where the best one is k=2.

Table 19. Table with results of valence for frequency domain, ternary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 93.5676 | 57.32484 | 58.04375 | 59.35401 | 58.52391 | 0.308738 | 0.032858 |
| SVM | 91.50685 | 60.50955 | 60.24115 | 62.22796 | 60.42343 | 0.014728 | 0.027881 |
| DT | 92.87671 | 52.86624 | 52.35583 | 55.87712 | 52.83629 | 0.012218 | 0.010025 |
| KNN | 77.53425 | 54.77707 | 57.50985 | 57.92133 | 54.03975 | 0.003622 | 0.012404 |

* Arousal results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 20. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 300 |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 1 |

Table 21. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 22.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 20 |
| min\_samples\_split: 2,5,10 | 5 |
| min\_samples\_leaf: 1,2,4 | 2 |



Figure . KNN performance for different values of k where the best one is k=2.

Table 23. Table with results of arousal for frequency domain, ternary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 98.898 | 86.38743 | 86.84361 | 87.1275 | 86.71717 | 0.771435 | 0.101065 |
| SVM | 81.57303 | 76.96335 | 77.03249 | 78.31029 | 76.4568 | 0.012616 | 0.034662 |
| DT | 96.62921 | 76.43979 | 76.13703 | 77.34255 | 76.37328 | 0.012246 | 0.010609 |
| KNN | 89.21348 | 78.53403 | 79.2731 | 79.06298 | 78.49042 | 0.004258 | 0.018876 |

* Dominance results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 24. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 200 |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 10 |
| min\_samples\_leaf: 1,2,4 | 1 |

Table 25. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.01 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 26.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 5 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 1 |

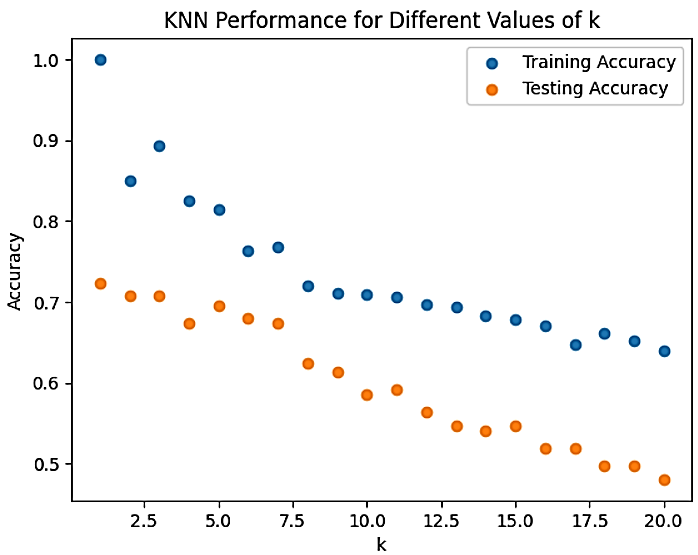


Figure 25. KNN performance for different values of k where the best one is k=3.

Table 27. Table with results of dominance for frequency domain, ternary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 97.39336 | 84.53039 | 85.63829 | 84.6687 | 84.67873 | 0.457282 | 0.066552 |
| SVM | 81.75355 | 71.8232 | 72.95804 | 72.82439 | 71.61278 | 0.012259 | 0.026511 |
| DT | 90.28436 | 81.21547 | 81.83889 | 81.81204 | 81.18587 | 0.008964 | 0.009735 |
| KNN | 89.33649 | 70.71823 | 70.28526 | 71.56863 | 70.26493 | 0.003747 | 0.014594 |

## Time Domain Results on Machine Learning Algorithms

In the time-domain analysis, we focused on extracting features directly from the raw EEG signals without employing any bandpass filtering. Instead, we computed statistical features such as the mean, standard deviation, skewness, and kurtosis for each EEG channel. These features were computed separately for baseline and stimulus data, and the resulting values were used to derive a ratio of stimulus to baseline, capturing the relative change in EEG characteristics.

To ensure the reliability of our models, we applied normalization to the time-domain features, standardizing them to have zero mean and unit variance using the “StandardScaler”. This step was crucial in ensuring that each feature contributed equally to the model's performance, avoiding biases due to different feature scales. The standardized features were then combined with participant-specific data such as age, gender, and subjective scores of valence, arousal, and dominance.

We used ternary and binary classification systems to group the emotional variables (dominance, arousal, and valence). The binary classification approach divided dominance, arousal, and valence into two categories: low and high. Binary conversion was performed on the given value, where 1 denotes a high level and 0 a low one. We also expanded our analysis to include ternary classification in order to encompass a more complex spectrum of emotional states. Three levels of classification were applied to the values, offering a deeper comprehension of the emotional states represented in the EEG signals.

We used random oversampling to reduce class imbalances in both binary and ternary classifications. In order to guarantee equal representation across all classes, this strategy duplicated instances from the minority class, improving the robustness and generalizability of our models. This oversampling strategy was essential for preventing model bias towards the majority class and ensuring equitable performance across all classes.

In order to classify based on the extracted and normalized time-domain features, we used a number of machine learning methods, such as Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), and K-Nearest Neighbor (K-NN). The models were assessed and trained using common metrics like F1-score, recall, accuracy, and precision. To further enhance model performance, hyperparameters for each model were optimized using grid search cross-validation as we have done for the frequency domain.

### Binary Classification

* Valance results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 28. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 100 |
| max\_depth: 5,10,20 | 5 |
| min\_samples\_split: 2,5,10 | 5 |
| min\_samples\_leaf: 1,2,4 | 2 |

Table 29. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 30.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 10 |
| min\_samples\_leaf: 1,2,4 | 1 |

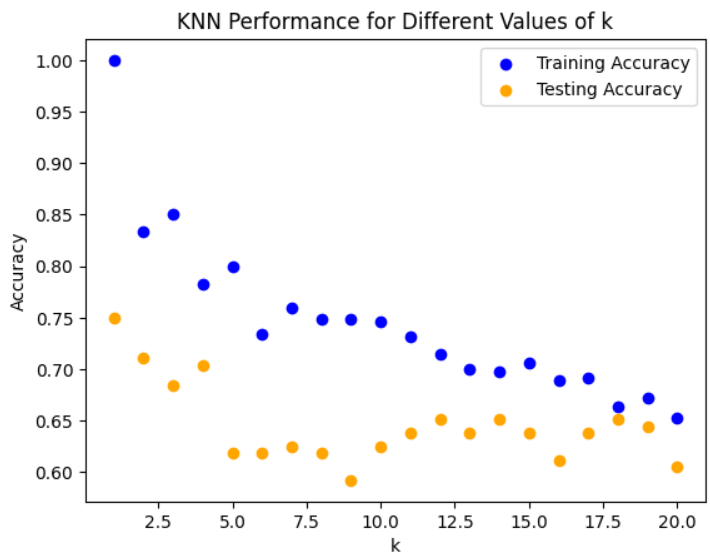


Figure 26. KNN performance for different values of k where the best one is k=2.

Table 31.Table with results of valance for time domain, binary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 94.35028 | 72.36842 | 72.61324 | 72.48787 | 72.34927 | 0.207545 | 0.039524 |
| SVM | 88.70056 | 71.05263 | 71.08014 | 70.96327 | 70.97222 | 0.010781 | 0.024876 |
| DT | 95.76271 | 65.13158 | 66.01513 | 65.40194 | 64.87465 | 0.01111 | 0.009527 |
| KNN | 83.33333 | 71.05263 | 77.05628 | 71.65627 | 69.71014 | 0.003841 | 0.012731 |

* Arousal results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 32. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 300 |
| max\_depth: 5,10,20 | 20 |
| min\_samples\_split: 2,5,10 | 5 |
| min\_samples\_leaf: 1,2,4 | 2 |

Table 33. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 34.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 20 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 1 |

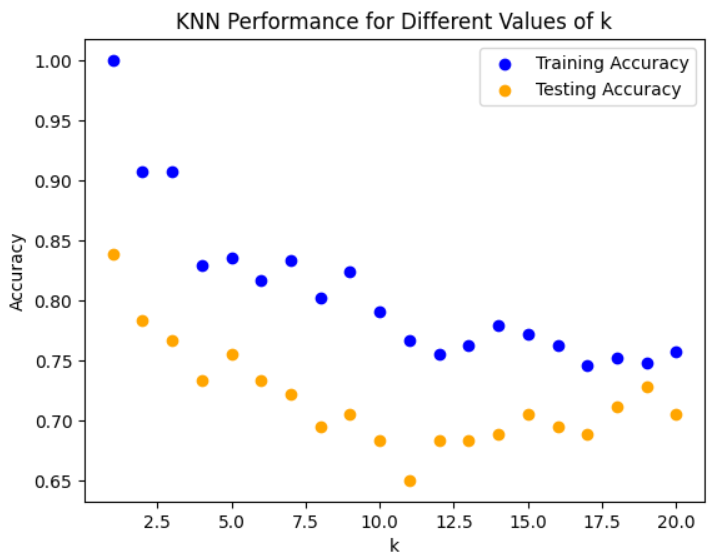


Figure 27. KNN performance for different values of k where the best one is k=2.

Table 35. Table with results of arousal for time domain, binary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 100 | 94.44444 | 94.83726 | 94.39437 | 94.42724 | 0.794008 | 0.088507 |
| SVM | 95.2381 | 87.77778 | 88.7619 | 87.68984 | 87.68044 | 0.010573 | 0.022135 |
| DT | 100 | 90 | 91.37114 | 89.89999 | 89.89899 | 0.014151 | 0.010448 |
| KNN | 90.71429 | 78.33333 | 82.20746 | 78.13928 | 77.57976 | 0.004632 | 0.014922 |

* Dominance results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 36. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 200 |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 5 |
| min\_samples\_leaf: 1,2,4 | 1 |

Table 37. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 38.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 20 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 2 |

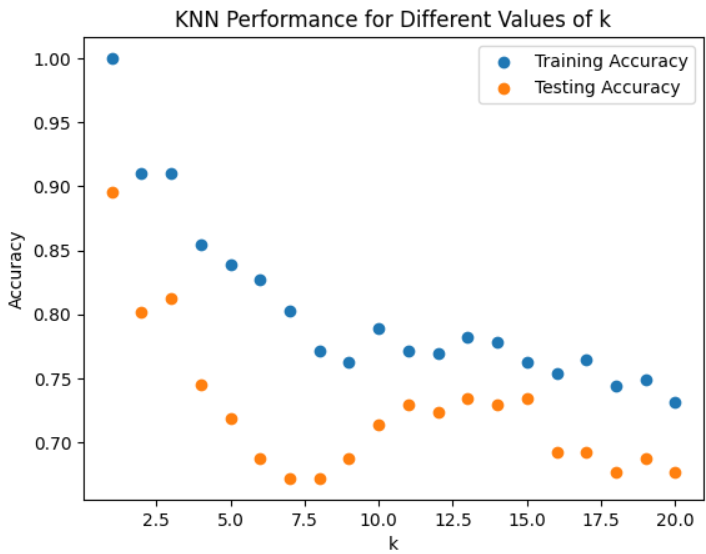


Figure . KNN performance for different values of k where the best one is k=2.

Table 39. Table with results of dominance for time domain, binary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 100 | 90.10417 | 90.39489 | 90.19757 | 90.09745 | 0.975223 | 0.10616 |
| SVM | 94.1704 | 81.25 | 82.32143 | 81.43726 | 81.14978 | 0.020211 | 0.039353 |
| DT | 97.53363 | 83.85417 | 84.42551 | 83.98828 | 83.81861 | 0.014769 | 0.014964 |
| KNN | 91.03139 | 80.20833 | 83.83838 | 80.54711 | 79.77827 | 0.008693 | 0.159092 |

### Ternary Classification

* Valence results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 40. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 100 |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 5 |
| min\_samples\_leaf: 1,2,4 | 1 |

Table 41. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 42.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 1 |

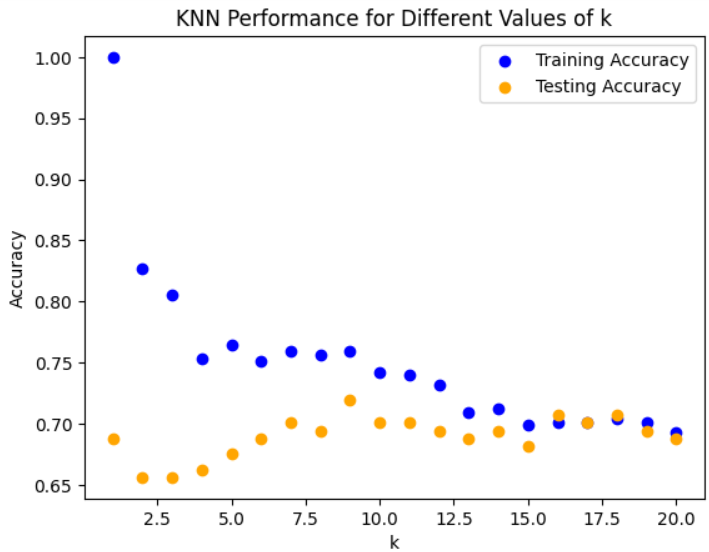


Figure . KNN performance for different values of k where the best one is k=2.

Table 43. Table with results of valence for time domain, ternary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 100 | 65.6051 | 65.38483 | 67.39305 | 66.12197 | 0.251198 | 0.048171 |
| SVM | 75.06849 | 67.51592 | 67.64024 | 68.42808 | 67.2911 | 0.013438 | 0.024571 |
| DT | 100 | 74.52229 | 74.5405 | 75.98048 | 74.30491 | 0.008985 | 0.009586 |
| KNN | 82.73973 | 65.6051 | 64.87179 | 67.80505 | 64.74305 | 0.003764 | 0.012791 |

* Arousal results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 44. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 200 |
| max\_depth: 5,10,20 | 20 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 1 |

Table 45. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel:Rbf, Linear, Poly | Rbf |

Table 46.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 20 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 2 |

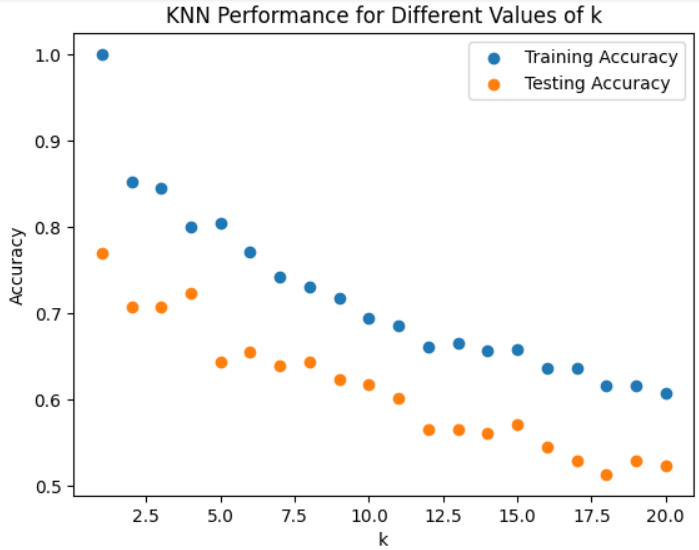


Figure . KNN performance for different values of k where the best one is k=2.

Table 47. Table with results of arousal for time domain, ternary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 100 | 86.91099 | 87.63577 | 87.72657 | 87.20375 | 0.611078 | 0.057867 |
| SVM | 97.52809 | 78.01047 | 77.99521 | 79.20123 | 78.00742 | 0.013202 | 0.026525 |
| DT | 97.30337 | 76.96335 | 76.61762 | 78.00307 | 76.81152 | 0.012586 | 0.009979 |
| KNN | 85.16854 | 70.68063 | 70.83054 | 71.36713 | 69.95413 | 0.003889 | 0.0263 |

* Dominance results:

Below we have the parameters for RF, SVM, DT and KNN that performed best for this experiment.

Table 48. Table that shows the best parameters of RF.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| n\_estimators: 100,200,300 | 200 |
| max\_depth: 5,10,20 | 10 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 2 |

Table 49. Table that shows the best parameters of SVM.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| C: 0.1, 1, 10 | 10 |
| Gamma: 0.1, 0.01, 0.001 | 0.1 |
| Kernel: Rbf, Linear, Poly | Rbf |

Table 50.Table that shows the best parameters of DT.

|  |  |
| --- | --- |
| Parameters | Best Parameters |
| max\_depth: 5,10,20 | 5 |
| min\_samples\_split: 2,5,10 | 2 |
| min\_samples\_leaf: 1,2,4 | 2 |

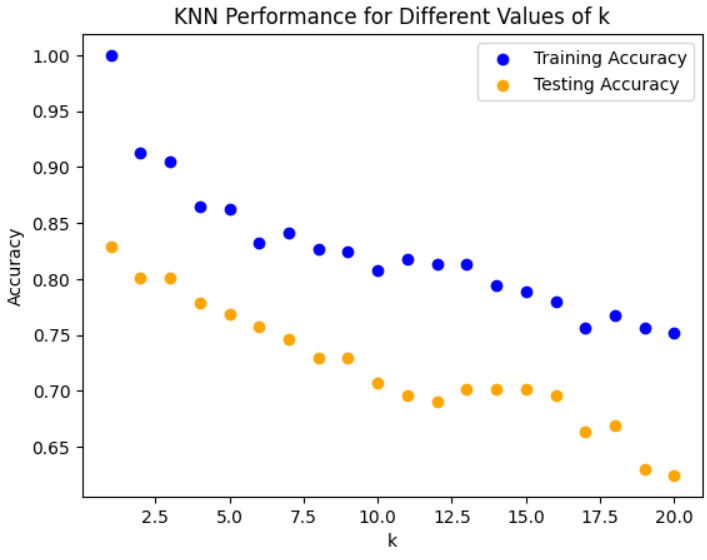


Figure . KNN performance for different values of k where the best one is k=3.

Table 51. Table with results of dominance for time domain, ternary classification using machine learning methods.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Training Accuracy | Test Accuracy | Precision | Recall | F1-score | Training Time (s) | Prediction Time (s) |
| RF | 99.76303 | 83.9779 | 84.92965 | 84.1785 | 84.228 | 2.390974 | 0.216562 |
| SVM | 87.91469 | 82.32044 | 82.61183 | 82.96146 | 82.3419 | 0.039497 | 0.084481 |
| DT | 89.57346 | 76.24309 | 76.22354 | 77.14672 | 76.01344 | 0.025087 | 0.029126 |
| KNN | 90.52133 | 80.1105 | 79.98994 | 80.83164 | 80.08304 | 0.017209 | 0.038126 |

## Frequency Domain Results on Deep Learning Algorithms

We improved upon the machine learning approach's feature extraction and normalization techniques in our deep learning analysis for the frequency domain. The main modifications were using deep learning models and modifying preprocessing stages to accommodate these models.

To ensure that the most relevant features were included in our dataset, we refined it through feature selection based on mutual information scores. The data was divided into training and testing sets, transformed into numpy arrays, and reshaped as necessary for deep learning models after class imbalances were addressed using RandomOverSampler.

Three deep learning models, a convolutional neural network (CNN), a deep neural network (DNN), and a recurrent neural network (RNN), were created and trained by us. Every model was created to efficiently handle the peculiarities of EEG data. The RNN employed LSTM layers to capture temporal relationships, the DNN used multiple Dense layers with dropout for regularization, and the CNN used Conv1D layers for feature extraction.

These models were put through standard criteria for training and evaluation, and the results were collected and stored for later study. Our goal was to enhance classification performance by identifying intricate patterns in EEG signals through the integration of deep learning.

### Binary Classification

* Valance results:

Table 52. Table with results of valence for frequency domain, binary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 68.96552 | 76 | 66.08696 | 70.69767 | 10.37528 |
| DNN | 70.93596 | 76.92308 | 69.56522 | 73.05936 | 2.418393 |
| RNN | 54.6798 | 100 | 20 | 33.33333 | 41.04441 |

* Arousal results:

Table 53. Table with results of arousal for frequency domain, binary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 69.0678 | 100 | 42.51969 | 59.66851 | 6.317367 |
| DNN | 76.27119 | 83.80952 | 69.29134 | 75.86207 | 3.856553 |
| RNN | 71.18644 | 74.78992 | 70.07874 | 72.35772 | 44.79735 |

* Dominance results:

Table 54. Table with results of dominance for frequency domain, binary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 89.58333 | 94.49541 | 84.42623 | 89.17749 | 12.9192 |
| DNN | 78.33333 | 76.1194 | 83.60656 | 79.6875 | 3.534346 |
| RNN | 77.5 | 76.98413 | 79.5082 | 78.22581 | 45.24107 |

### Ternary Classification

* Valance results:

Table 55. Table with results of valence for frequency domain, ternary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 45.22293 | 46.35687 | 45.22293 | 43.9447 | 3.747469 |
| DNN | 39.49045 | 39.48135 | 39.49045 | 37.20828 | 4.096674 |
| RNN | 41.40127 | 41.06145 | 41.40127 | 40.46282 | 28.32695 |

* Arousal results:

Table 56. Table with results of arousal for frequency domain, ternary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 73.82199 | 74.56444 | 73.82199 | 74.07942 | 3.808491 |
| DNN | 63.35079 | 64.94828 | 63.35079 | 55.01118 | 3.640051 |
| RNN | 30.89005 | 9.541953 | 30.89005 | 14.5801 | 34.71046 |

* Dominance results:

Table 57. Table with results of dominance for frequency domain, ternary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 62.43094 | 68.67216 | 62.43094 | 61.73756 | 11.32303 |
| DNN | 58.56354 | 58.31493 | 58.56354 | 57.33642 | 7.121601 |
| RNN | 63.53591 | 62.60239 | 63.53591 | 62.72972 | 45.79808 |

## Time Domain Results on Deep Learning Algorithms

We modified our preprocessing and feature extraction procedures from the machine learning technique for the deep learning time-domain analysis. From the raw EEG data, we retrieved characteristics such as mean, standard deviation, skewness, and kurtosis. Mutual information scores were utilized for feature selection, and RandomOverSampler was employed to rectify any class imbalances.

Our frequency-domain analysis was comparable to the deep learning implementation. To categorize the time-domain features, we used recurrent neural networks (RNNs), deep neural networks (DNNs), and convolutional neural networks (CNNs). These models were trained and assessed with the use of measures like F1 score, accuracy, precision, and recall. This method improved the classification performance for EEG time-domain signals by utilizing neural networks' sophisticated pattern recognition capabilities.

### Binary Classification

* Valance results:

Table 58. Table with results of valence for time domain, binary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 62.56158 | 75.32468 | 50.43478 | 60.41667 | 10.44621 |
| DNN | 60.09852 | 71.79487 | 48.69565 | 58.03109 | 4.501724 |
| RNN | 43.84236 | 100 | 0.869565 | 1.724138 | 25.73453 |

* Arousal results:

Table 59. Table with results of arousal for time domain, binary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 81.77966 | 89.62264 | 74.80315 | 81.54506 | 4.174967 |
| DNN | 72.0339 | 92.95775 | 51.9685 | 66.66667 | 6.632348 |
| RNN | 46.61017 | 100 | 0.787402 | 1.5625 | 45.20587 |

* Dominance results:

Table 60. Table with results of dominance for time domain, binary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 90 | 89.51613 | 90.98361 | 90.2439 | 16.09212 |
| DNN | 84.16667 | 100 | 68.85246 | 81.5534 | 15.57864 |
| RNN | 76.25 | 69.23077 | 95.90164 | 80.41237 | 52.81654 |

### Ternary Classification

* Valence results:

Table 61. Table with results of valence for time domain, ternary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 47.7707 | 49.90671 | 47.7707 | 45.9676 | 3.584791 |
| DNN | 40.76433 | 48.00371 | 40.76433 | 36.13773 | 3.404319 |
| RNN | 27.38854 | 7.501319 | 27.38854 | 11.77707 | 21.59148 |

* Arousal results:

Table 62. Table with results of arousal for time domain, ternary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 63.87435 | 63.46216 | 63.87435 | 60.44466 | 9.855819 |
| DNN | 54.97382 | 52.60231 | 54.97382 | 48.0975 | 12.63747 |
| RNN | 30.89005 | 9.541953 | 30.89005 | 14.5801 | 46.10062 |

* Dominance results:

Table 63. Table with results of dominance for time domain, ternary classification using deep learning methods.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Test Accuracy | Precision | Recall | F1 Score | Training Time (seconds) |
| CNN | 72.37569 | 72.63271 | 72.37569 | 72.31866 | 7.987695 |
| DNN | 63.53591 | 63.99848 | 63.53591 | 61.45505 | 3.511997 |
| RNN | 30.93923 | 41.38091 | 30.93923 | 15.37061 | 28.41542 |

**CHAPTER 4**

# CONCLUSION

This thesis has thoroughly investigated the use of deep learning and machine learning methods to identify emotions from EEG signals from the DREAMER dataset. The study assessed the effectiveness of different classification methods in the frequency and temporal domains using thorough preprocessing and feature extraction. The results show that the frequency domain features are robust and effective in capturing the subtleties of EEG signals, as they consistently produced superior results across various classification tests. Notably, especially when applied to deep learning models, frequency domain-based techniques fared better than time domain methods in both binary and ternary classifications.

But the study also showed that there was an anomaly in the ternary classification machine learning paradigms, where the frequency domain features did not perform as well as expected. This emphasizes how difficult and varied emotion recognition tasks may be, implying that there isn't a particular domain or approach that works best for everyone. This work opens new avenues for EEG-based emotion recognition research, with possible applications in the fields of affective computing, healthcare, and human-computer interaction. By investigating hybrid models that use both frequency and time domain information or by utilizing more complex deep learning architectures to improve classification accuracy, future research could build on these findings.

# 

# REFERENCES

1. Kamble K, Sengupta J. A comprehensive survey on emotion recognition based on electroencephalograph (EEG) signals. Multimed Tools Appl. 2023;82(18):27269–304.

2. Plutchik R. A general psychoevolutionary theory of emotion. In: Theories of emotion. Elsevier; 1980. p. 3–33.

3. s-l1600.png (1003×531) [Internet]. [cited 2024 May 22]. Available from: https://i.ebayimg.com/images/g/VlcAAOSwwadiBdbs/s-l1600.png

4. Vempati R, Sharma LD. A systematic review on automated human emotion recognition using electroencephalogram signals and artificial intelligence. Results in Engineering. 2023;101027.

5. The main parts of the human brain indicate the Frontal, Parietal,... | Download Scientific Diagram [Internet]. [cited 2024 May 22]. Available from: https://www.researchgate.net/figure/The-main-parts-of-the-human-brain-indicate-the-Frontal-Parietal-Temporal-and-Occipital\_fig7\_352676025

6. Hairston WD, Whitaker KW, Ries AJ, Vettel JM, Bradford JC, Kerick SE, et al. Usability of four commercially-oriented EEG systems. J Neural Eng. 2014;11(4):046018.

7. Sohaib AT, Qureshi S. An empirical study of machine learning techniques for classifying emotional states from EEG data. 2012.

8. A person with Electrodes attached to the sensor | Download Scientific Diagram [Internet]. [cited 2024 May 22]. Available from: https://www.researchgate.net/figure/A-person-with-Electrodes-attached-to-the-sensor\_fig3\_363200648

9. Support Vector Machine (SVM) Algorithm - Javatpoint [Internet]. [cited 2024 May 22]. Available from: https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm

10. Rigatti SJ. Random forest. J Insur Med. 2017;47(1):31–9.

11. House Price Prediction in Python using Random Forest | by Alexander Rubino | Medium [Internet]. [cited 2024 May 22]. Available from: https://medium.com/@alexrubino05/house-price-prediction-in-python-using-random-forest-59dc51bf7498

12. Patel HH, Prajapati P. Study and analysis of decision tree based classification algorithms. International Journal of Computer Sciences and Engineering. 2018;6(10):74–8.

13. CART (Classification And Regression Tree) in Machine Learning - GeeksforGeeks [Internet]. [cited 2024 May 22]. Available from: https://www.geeksforgeeks.org/cart-classification-and-regression-tree-in-machine-learning/

14. Most Popular Distance Metrics Used in KNN and When to Use Them - KDnuggets [Internet]. [cited 2024 May 22]. Available from: https://www.kdnuggets.com/2020/11/most-popular-distance-metrics-knn.html

15. Deng L, Yu D. Deep learning: methods and applications. Foundations and trends® in signal processing. 2014;7(3–4):197–387.

16. (5) What is Convolutional Neural Network — CNN (Deep Learning) | LinkedIn [Internet]. [cited 2024 May 22]. Available from: https://www.linkedin.com/pulse/what-convolutional-neural-network-cnn-deep-learning-nafiz-shahriar/

17. An Overview of Deep Belief Network (DBN) in Deep Learning [Internet]. [cited 2024 May 22]. Available from: https://www.analyticsvidhya.com/blog/2022/03/an-overview-of-deep-belief-network-dbn-in-deep-learning/

18. Introduction to FeedForward Neural Networks | by Yash Upadhyay | Towards Data Science [Internet]. [cited 2024 May 22]. Available from: https://towardsdatascience.com/feed-forward-neural-networks-c503faa46620

19. Satapathy SK, Dehuri S, Jagadev AK, Mishra S. EEG brain signal classification for epileptic seizure disorder detection. Academic Press; 2019.

20. Probabilistic Neural Networks: An Introduction to Probability Theory in Machine Learning | by Kajeeth Kumar G | Artificial Intelligence in Plain English [Internet]. [cited 2024 May 22]. Available from: https://ai.plainenglish.io/probabilistic-neural-networks-an-introduction-to-probability-theory-in-machine-learning-844091af3500

21. Zangeneh Soroush M, Maghooli K, Kamaledin Setarehdan S, Nasrabadi AM. A review on EEG signals based emotion recognition. International Clinical Neuroscience Journal. 2017;4(4):118–29.

22. Sokolova M, Japkowicz N, Szpakowicz S. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. In: Australasian joint conference on artificial intelligence. Springer; 2006. p. 1015–21.

23. Katsigiannis S, Ramzan N. DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. IEEE J Biomed Health Inform. 2017;22(1):98–107.

24. Alasadi SA, Bhaya WS. Review of data preprocessing techniques in data mining. Journal of Engineering and Applied Sciences. 2017;12(16):4102–7.