Chapter 1

Introduction

1.1. Brain-Computer Interface (BCI)

Brain-Computer Interface **(**BCI) system can be defined as a communication scheme that facilitates the human brain to communicate with external intelligent devices such as an electronic mouse, computers, smart washing machines, neuro-prosthetics, visual simulation of games, robotic components (arms, fingers...), and medical assistive devices [1, 2, 3]. BCI systems acquire task-specific electrical signals from the human brain and analyze and translate them into computer-based control commands using multiple signal processing methods. In general, the translated commands are used to restore lost neurological functions caused by stroke or any nervous system injuries [4]. The primary objective of the newly developed BCI system is to improve communication ability between physically disabled persons and different prosthetic devices such as incontinence control machines, hearing aids, and robotic limbs.

In BCI communication, Motor Imagery (MI) is an important paradigm that deals with imagination-based muscular activities. A subject is required to imagine in his brain about a limb's movement without performing any physical task [5]. Simultaneously, the brain emits the signals corresponding to imagined limb activity and these oscillations may be recorded through multiple channels or electrodes. In a multichannel EEG setup, several electrodes, usually ranging from a few to several dozen, are strategically positioned across different regions of the scalp. Each electrode acts as a sensor that picks up the electrical fluctuations produced by groups of neurons firing in synchronization. By placing electrodes in different locations, researchers can capture brain activity from various brain regions simultaneously. This approach facilitates understanding the dynamics of the complex brain activities during MI activities' execution.

The BCI system aims to detect and quantify features of recorded brain signals that indicate the user's intentions and to translate these features in real-time into device commands that accomplish the user's intent. To achieve this, it consists of four sequential steps: (1) signal acquisition, (2) feature extraction, (3) feature selection, and (4) classification. These phases are controlled by an operating protocol that defines the onset and timing of the operation, the details of signal processing, the nature of the device commands, and the oversight of performance.

However, the implementation of MI-based BCI systems is notoriously difficult because their training procedure is time-consuming and shows limited BCI channel capacity. The block diagram of a generalized BCI framework is shown in Figure 1.1.

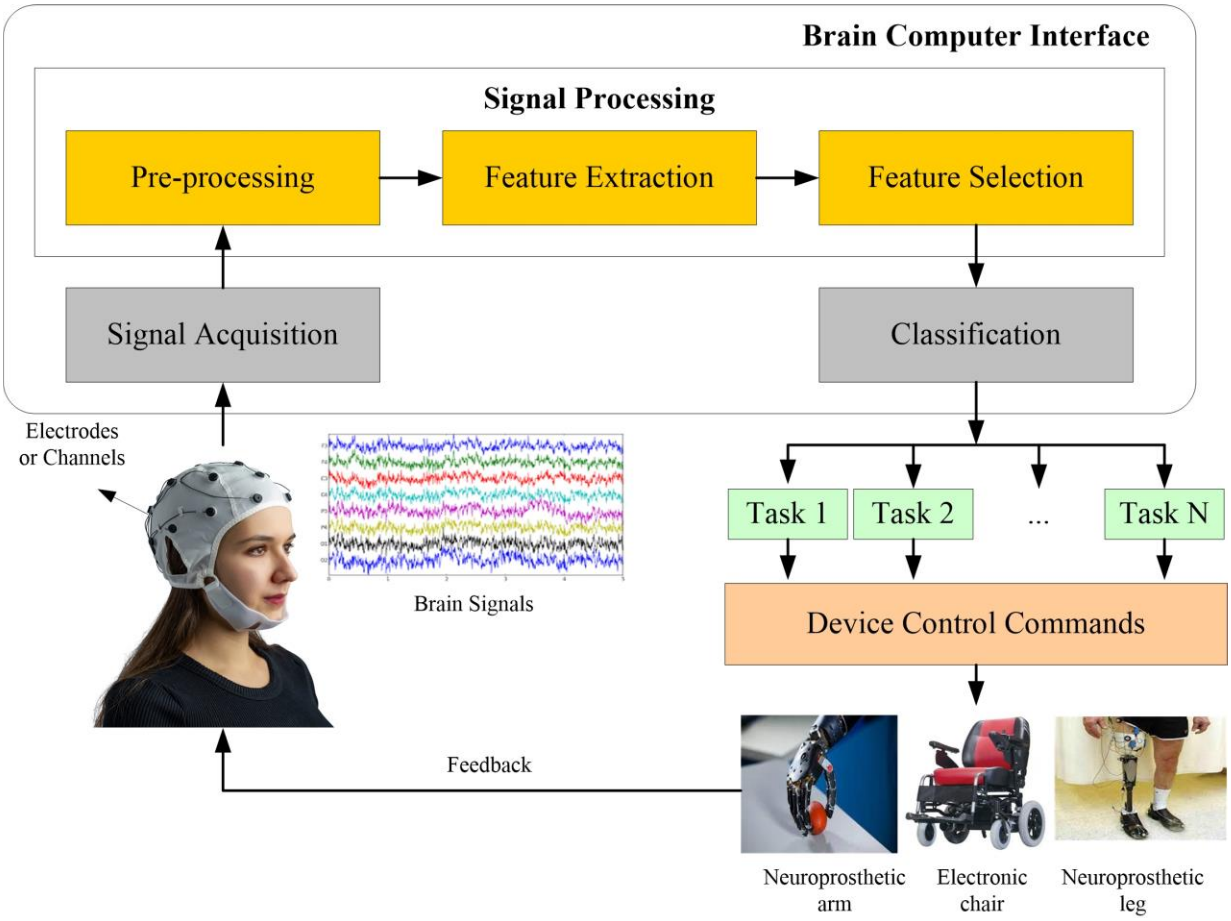

Figure 1.1. A generalized block diagram of the Brain-Computer Interface (BCI)

1.2. Electroencephalography in BCI

BCI systems take advantage of various neuroimaging techniques such as Electroencephalography (EEG), Magnetoencephalography (MEG), functional Magnetic

Resonance Imaging (fMRI), and functional Near-Infrared Spectroscopy (fNIRS) are frequently used to collect signals from the human brain [6]. Although some noninvasive technologies provide a higher spatial resolution (e.g. fMRI), the EEG has proved to be the most popular method due to its direct measures of neural activity, lower cost, and portability for clinical use [7]. EEG measures electrical brain activity caused by the flow of electric currents during synaptic excitations of neuronal dendrites, especially in the cortex, but also in the deep brain structures. In general, EEG signals are non-stationary, chaotic, and multichannel in nature. Therefore, it is challenging to identify true cognitive patterns associated with the acquired motor-imagery-specific EEG signals. As a result, the investigation of relationships between EEG signals and upper limb movement, real and imaginary, has become a fascinating area of research in recent years.

Several researchers defined the mapping between motor movement and the corresponding mental states by discriminating among brain oscillations [8, 9]. Researchers have classified these oscillations into four different classes: (1) delta waves (below 4 Hz), (2) theta waves (frequency in the range of 4.0–8.0 Hz), (3) alpha waves (frequency in the range of 8.0–13.0 Hz) and (4) beta waves (≥ 13.0 Hz). They concluded that a complex EEG spectrum of the upper range of alpha waves and a lower range of beta waves is responsible for most individuals' muscular movement and related activities. In addition, they concluded that only a small segment of the brain, namely the Motor Cortex (MC) is mainly responsible for executing MI activities and generating Beta waves. In Figure 1.2, we have shown the spatial location of MC and other cognitive regions in the human brain. Moreover, the role of other brain regions in MI execution is still an open question for new researchers.

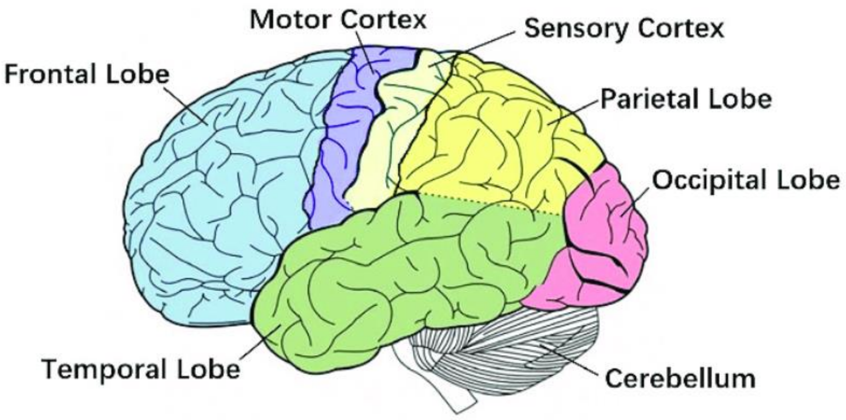

Figure 1.2. The spatial location of the Motor Cortex (in dark blue color) region in the brain

1.3. Types of BCI Systems

Based on signal collection techniques, BCI systems can be categorized into three groups:

1.3.1. Invasive BCI Systems

In invasive BCI systems, the array of microelectrodes is implanted into a specific cortex during neurosurgery to measure electrical potentials produced by activated neurons. Invasive types of BCI are implanted directly into the brain during neurosurgery. There are single-unit BCIs, that detect signals from a single area of brain cells, and multiunit BCIs that detect signals from multiple areas. Best-quality signals are recorded in this method, but it also involves various problems, such as the risk of forming scar tissues. The body reacts to the foreign object and builds a scar around the electrodes, which causes deterioration of the quality of signals. Because neurosurgery can be a risky and expensive process, the target of invasive BCI is mainly blind and paralyzed patients.

1.3.2. Semi-invasive BCI Systems

In semi-invasive BCI systems, electrodes are placed on the exposed surface of the brain to measure electrical activity from the cerebral cortex. It was used for the first time in the 1950s at the Montreal Neurological Institute. It is called semi-invasive but it still requires a craniotomy to implant the electrodes. For this reason, it is used only when surgery is necessary for medical reasons (epilepsy for example). The electrodes may be placed outside the dura mater (epidural) or under the dura mater (subdural). The strip or grid electrodes cover a large area of the cortex (from 4 to 256 electrodes), allowing a diverse range of cognitive studies.

1.3.3. Non-invasive BCI Systems

Noninvasive EEG-based BCIs are the most widely explored approach owing to the minimal risk involved and the relative convenience of conducting studies and recruiting participants. The applications to date are generally limited to low-degree-of-freedom continuous movement control and discrete selection. Most non-invasive BCI systems use electroencephalogram (EEG) signals; i.e., the electrical brain activity recorded from electrodes placed on the scalp. The main source of the EEG is the synchronous activity of thousands of cortical neurons. Measuring the EEG is a simple noninvasive way to monitor electrical brain activity, but it does not provide detailed information on the activity of single neurons (or small brain areas). These signals efficiently record brain activity as it unfolds in real-time, at the level of milliseconds (thousandths of a second).

Considering the broader scope of non-invasive BCI systems over the remaining two, our study mainly focuses on the behavioral analysis of EEG signals in the context of non-invasive MI-based BCI systems.

1.4. Major Issues in BCI Design

In the last two decades, substantial BCI innovations have been performed to develop robust BCI systems and related assistive technologies. However, there are many crucial issues and challenges in different BCI components that degrade the machine's overall performance when used with a real-time application. The BCI community should address these issues to make further advances in BCIs. A short discussion on prominent BCI design issues is given in subsequent subsections.

1.4.1. Issues Related to EEG Modality in BCI Design

It is known that EEG is one of the BCI system's preferred modalities for signal acquisition. In general, EEG oscillations are highly chaotic, non-stationary, multichannel, and have poor signal-to-noise (SNR) ratio [10]. Therefore, a series of preprocessing steps are required to improve the signal quality in terms of reducing unwanted peaks and improving the signal's strength. Good preprocessing leads to achieving better feature separation and classification performance in the discrimination of multiple MI tasks. Although a variety of methods such as signal filtering, signal transformations, regressions, and blind source separation have been proposed, they have their limitations. For example, filtering methods like low, high, and band-pass filters perform well when the frequency bands of the signals do not overlap. The performance of alternative methods such as Independent Component Analysis (ICA), Principal Component Analysis (PCA), and Canonical Correlation Analysis (CCA) also depends on the types of noise present in extracted signals [11, 12]. Based on an individual's limitations, it is required to demonstrate the suitability of the applied preprocessing methods with the extracted EEG signals.

Determining the suitable feature extraction method and precise classification approach is another challenge in the discrimination of cognitive patterns associated with extracted EEG signals. An inappropriate data modeling strategy may reduce the accuracy and robustness of the implemented BCI classification model. In addition, such methods may suffer from overfitting issues which limits the generalizability of the model when used with unseen data. Overfitting of the model causes it to perform perfectly on the training set while fitting poorly on the testing set. This is because the over-fitted model has difficulty coping with pieces of the information in the testing set, which may be different from those in the training set. On the other hand, over-fitted models tend to memorize all the data, including unavoidable noise on the training set, instead of learning the discipline hidden behind the data. Therefore, strong sequential steps are required to model EEG signals for BCI system development.

1.4.2. Channel Redundancy

It has been discussed that EEG signals are multichannel in nature. In EEG signal processing, MI-specific signals are collected from definite scalp sites referred to as channels or electrodes. Although the dense arrangement of electrodes on the human scalp explores more information about the underlying mechanism of neuronal activity but induces an abundant amount of redundant and irrelevant information in the form of noise and high-dimensional data. It has increased the requirement of an effective channel selection mechanism threefold: (1) To reduce the computational complexity of an applied MI state prediction model by selecting specific sets from EEG signals, and (2) To minimize the amount of overfitting that redundant channels may cause, and (3) To reduce the time complexity of the complete BCI setup in different applications. Keeping these three requirements in mind, developing a better solution to find the optimal channel set in the BCI development process is required.

1.4.3. Issues with Performance Evaluation Metrics

In previous studies, several performance metrics have been developed to investigate the accuracy of the applied signal-processing method. However, it is almost impossible to compare the same type of BCI systems when dissimilar performance metrics evaluate them. Hence, the BCI research community should recommend a standard and systematic approach or a single metric to quantify a specific BCI application. For example, the number of control commands, types of control commands, distance covered, the time required, number of collisions, classification accuracy, average success rate, and Channel Reduction Rate (CRR) should be utilized to evaluate the performance of BCI wheelchair control. However, the aforesaid metrics cannot be used to evaluate BCI-controlled video games Therefore, it is required to induce a set of standard performance measures for a specific group of BCI systems, that can help researchers validate and compare their results with other state-of-the-art methods.

1.4.4. Curse of Dimensionality

In the growing era of the information industry, data size and features have enormously increased, whereas the patterns are relatively few, which results in relevant, redundant, and irrelevant information about the features. Such redundant and irrelevant features increase the computational burden and bring the “Curse of Dimensionality” (CoD), which refers to those phenomena that arise during data analysis in high dimensional space [13, 14] The common theme of these problems is that when the dimensionality increases, the volume of the space increases so fast that the available data becomes sparse. In machine learning, sparse data may alter the performance of applied optimization algorithms and their ability to calculate accurate predictions. In addition, sparse data increases the spatial and temporal complexity of machine learning algorithms because very large matrices require a lot of memory and abundant mathematical operations to complete the desired tasks.

Even with the steady increase in the speed of modern computers and the availability of cheaper parallel and cloud computing facilities, this does not ease the challenges of high-dimensional information analysis. Efforts to develop new methods and tools are still highly needed. It may need a paradigm shift and a nonconventional way of thinking to problem-solve concerning high-dimensional data. In addition, noise and uncertainties often exist in high-dimensional datasets. Such noisy data can become more challenging to process and may thwart the proper application of any data mining technique. Therefore, algorithms tend to be problem-specific and even data-specific. Thus, there is no generic approach in general.

1.5. Scope of Combinatorial Optimization in BCI

The proposed study attempts to determine the most reliable and accurate solutions for all four issues mentioned above by recognizing the true data patterns. Combinatorial optimization is a subfield of mathematical optimization that consists of finding an optimal solution from a finite set of solutions, where the set of feasible solutions is always discrete. In many such problems, where the solution space is very large, conventional exhaustive search methods are not tractable, and so specialized algorithms that quickly rule out large parts of the search space or approximation algorithms must be resorted to, instead. It is known that combinatorial search algorithms are typically concerned with problems that are NP-hard [15]. However, generic search algorithms are not guaranteed to find an optimal solution first, nor are they guaranteed to run quickly (in polynomial time). Combinatorial optimization problems can be viewed as searching for the best element from a set of discrete items; therefore, in principle, any sort of search algorithm or metaheuristic can be used to solve them. It gives them an edge over traditional brute-force methods to solve the above-discussed problems.

In our thesis, we considered the discussed issues as a set of discrete problems therefore, multiple combinatorial optimization algorithms are successfully implemented to rule out the large part of the solution space and get the optimal solution for all four problems.

1.6. Thesis Objectives

This thesis work targets five objectives that pertain to the above-mentioned five critical issues in BCI design.

1.6.1. Mapping EEG Signals to Motor Activities

This objective refers to a classification process where an applied machine learning technique recognizes the performed task by understanding the structural and functional patterns in respective EEG signals. To achieve this objective, a subject-specific classification model is developed where an adaptive classifier is trained subject by subject and its performance is evaluated on test data. This objective focuses on the analysis of existing signal pre-processing methods, feature extraction, and classification algorithms followed by the design of a BCI framework. It includes a detailed study of the following computational steps:

1.6.1.1. Collecting Signals Related to Motor Imagery and Signal Preprocessing

The experiments are carried out on three public datasets of EEG-based motor activities provided by BCI communities. Then, the acquired raw EEG signals are preprocessed using different filtering and feature extraction techniques. In our demonstrations, two signal preprocessing techniques: (1) Savitzky-Golay, and (2) Butterworth band pass filtering methods are employed to improve the signal quality and minimize the effect of noise and artifacts.

1.6.1.2. Feature Engineering

In this process, each preprocessed EEG signal is represented by a set of parameters (features). Then, Feature Normalization is performed to avoid possible problems caused by inadequately scaled features. This part of the EEG processing is crucial because it provides the ability to distinguish between different classes. It thus directly affects the accuracy of the final classification. In this step, two feature extraction techniques: (1) Multivariate Empirical Mode Decomposition (MEMD), and (2) Regularized Common Spatial Pattern (RCSP) are used to extract MI-specific features from given EEG signals.

1.6.1.3. MI-Tasks Classification Using Machine Learning Methods

Various supervised and unsupervised classification methods can be used to obtain the final results of the analysis. In our experiments, a set of different discriminating algorithms: (1) eXtreme Gradient Boosting (XGB), (2) Support Vector Machine and its Regularized variant (SVM and RSVM), (3) Naïve Bayes (NB), and (4) Decision Tree (DT) are employed for the detection of true motor imagery tasks. Moreover, the suitability of applied classifier techniques directly depends on the maximum classification accuracy that is realized after discriminating the motor-imagery-specific mental-state patterns.

1.6.2. Optimal Channel Selection in Brain-Computer Interface Development

In EEG signal processing, MI-specific signals are collected from definite scalp sites referred to as channels or electrodes (Multichannel behaviour). In this thesis, multichannel EEG signal processing for MI task discrimination is addressed as a multiclass classification problem. Although the dense arrangement of electrodes on the human scalp explores more information about the underlying mechanism of neuronal activity but induces an abundant amount of redundant and irrelevant information in the form of noise and high-dimensional data. In addition, it increases the hardware complexity of the BCI system, which requires more effort during the BCI preparation setup. Therefore, it is essential to reduce these efforts using a minimal but informative set of channels. This challenge is also termed Optimal Channel Selection (OCS) in designing a BCI system [16].

1.6.3. Finding the Role of Cognitive Regions in Motor-Imagery Activities

The brain is the most complex organ in the human body. It produces our every thought, action, memory, feeling, and experience of the world. The anatomical structure of the human brain with seven different cognitive regions is shown in Figure 1.2. The list of cognitive regions are: (1) Frontal lobe, (2) Motor cortex, (3) Sensory cortex, (4) Parietal lobe, (5) Occipital lobe, (6) Cerebellum, and (7) Temporal lobe. Each one of them has its spatial location in the human brain and functions to execute a specific set of tasks. For example, the frontal lobe is important for voluntary movement, expressive language, and managing higher-level executive functions. Similarly, the motor cortex is mainly involved in the planning, control, and execution of voluntary movements. The functional properties of other regions are explicitly given in Table

1.1. However, a comprehensive formalization of each region's function with the motor imagery task is still an open question for new researchers. In this thesis, we attempt to determine the role of different cognitive regions excluding the motor cortex in the execution of limb or voluntary movements.

Table 1.1. Functions of different cognitive signal bands

Sr. No.	Signal band	Functions
1	Gamma (> 30Hz)	Learning, memory, and information processing (Decision making)
2	Beta (13- 30Hz)	Conscious thought and logical thinking, muscular stimulation effect (Motor Imagery activities)
3	Alpha (8-12Hz)	Calmness, alertness of mental state, mind and body integration, and learning
4	Theta (4-8Hz)	Sleeping or dreaming, but they don't occur during the deepest phases of sleep, "drowsiness" i.e. the first stage of sleep
5	Delta (< 4Hz)	Deep sleep

1.6.4. Maximizing the Performance of the BCI Systems

In general, many factors influence BCI performance; considering the underlying cortical-subcortical networks is of crucial importance. For example, MI-induced signals are best recorded from the motor cortex and associated neighborhood regions. Therefore, an accurate neurophysiological basis of the human brain is required to capture the true motor-imagery signals that further help to improve the global performance of a generalized BCI system. Another, major factor such as maintaining an acceptable SNR in the non-invasive long-term recording is critical and vitally influences the classification accuracy. These signals are highly susceptible to different types of noises that cause short and long-term signal variation within and across individuals. Due to these intrinsic signal variations, BCI systems require subject-specific training, during which a subject attends a calibration session that is tedious and often frustrating. Therefore, minimizing the average calibration time also leads to boosting the real-time performance of BCI performance. Other components such as classification accuracy, minimum hardware setup, and suitable data analysis methods also influence the performance of the BCI framework. Considering the mentioned factors, this thesis concentrates on improving the overall performance of a given BCI system in terms of high classification accuracy, light hardware components, and less calibration time.

1.6.5. Dealing with High-Dimensional Data Using Feature Selection

High dimensional data refers to a dataset in which the number of features (*p*) is larger than the number of observations (*n*) [17]. Feature selection is important because irrelevant and redundant features induce a lot of noise that reduces the overall classification accuracy. To solve this problem, the proposed thesis aims to determine the most feasible solution by eliminating irrelevant and redundant alternatives from a large hyperspace. The emergence of metaheuristics for solving such optimization problems is one of the most notable achievements of the last two decades in operations research. These algorithms are imperative to find a near-optimal solution based on imperfect or incomplete information in this real world of limited resources (e.g., computational power and time). Thus, metaheuristics can often find good solutions with less computational effort than optimization algorithms, iterative methods, and simple greedy heuristics.

In this objective, we assume that metaheuristics are one of the powerful optimization techniques that have better search-space exploration ability than the conventional feature selection methods [18]. In addition, these methods outperform other methods on high-dimensional datasets when merged with a suitable filtering criterion [19]. In light of the aforementioned facts, this thesis merges the instance update strategy of the recently introduced metaheuristic algorithm, Dynamic Butterfly Optimization Algorithm (DBOA) with an interaction maximization scheme and determines the optimal feature subset. This problem is similar to the Optimal Channel Subset (OCS) selection problem which seeks the most optimal and correlated channels from the raw channel set. We validated our proposed feature selection method on twenty publicly open datasets using three classifiers: (1) Support Vector Machine (SVM), (2) Naïve Bayes (NB) algorithm, and (3) Decision Tree (DT). Also, a five-fold cross-validation approach is used to quantify classification results statistically to split the global population into training and evaluation sets.

1.7. Performance Measures Used in the Study

Choosing the right metric is crucial while evaluating the performance of the applied machine learning (ML) models. In some applications, the application of a single metric may not provide a detailed view of the used computational steps therefore, it is required to use a subset of the independent metrics to have a concrete evaluation of the used mathematical concepts. In this thesis, we consider different performance metrics to measure and compare the performance of our methods. The details of the used measures are discussed below:

1.7.1. Average Classification Accuracy

Average Classification accuracy (ACA) describes a classifier's ability to discriminate between samples using a selected optimal channel set [20]. The classification accuracy can be calculated using Eq. 1.1.

$$Average\ Classification\ Accuracy\ (ACA) = \frac{1}{M}\sum_{i=1}^{M}\frac{1}{N}\sum_{j=1}^{N} match\ (C_j, L_j) \quad (1.1)$$

where M is the total number of iterations the algorithm has executed, N represents the total number of observations in the test dataset, C_j and L_j indicate predicted and true class labels, respectively, and $match$ is a comparison function that provides output 1 when both labels are the same and outputs 0 when both are different.

1.7.2. Channel/ Feature Reduction Rate

Channel / Feature Reduction Rate [21] is an indicator that shows how effectively the relevant channels or features are selected. It is an important measure that reduces the BCI framework's hardware complexity without compromising the classification accuracy. The CRR in this study is computed using Eq. 1.2.

$$Channel\ Reduction\ Rate\ (CRR) = 1 - \frac{Number\ of\ selected\ channels}{Total\ number\ of\ channels} \quad (1.2)$$

1.7.3. Confusion Matrix

Confusion or error matrix is one of the key concepts in classification analysis, which is a tabular visualization of the model predictions versus the ground-truth labels [22]. Each row of the confusion matrix represents the instances in a predicted class and each column represents the instances in an actual class. Each row of the matrix represents the instances in an actual class while each column represents the instances in a predicted class. The template for any binary confusion matrix uses the four kinds of results (true positives, false negatives, false positives, and true negatives) along with the positive and negative classifications. This allows for a more detailed analysis than simply observing the proportion of correct classifications (accuracy). Accuracy will yield misleading results if the data set is unbalanced; that is when the numbers of observations in different classes vary greatly. The four outcomes can be formulated in a 2×2 confusion matrix, as follows:

Predicted Condition

Actual Condition

Total Population (P+N)	Positive (PP)	Negative (PN)
Positive (P)	True positive (TP)	False negative (FN)
Negative (N)	False positive (FP)	True negative (TN)

Based on the above-mentioned confusion matrix, the following measures are also evaluated:

1.7.4. Precision

Precision is the ratio between the True Positives and all the Positives. It is defined in Eq. 1.3.

$$\textbf{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{1.3}$$

1.7.5. Recall

Recall is the measure of our model correctly identifying True Positives and it is defined in Eq. 1.4.

$$\textbf{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{1.4}$$

1.7.6. F1-Score

This metric conveys the balance between precision and recall. It can be calculated using Eq. 1.5.

$$\textbf{F1-Score} = \frac{2 * Precision * Recall}{Precision + Recall} \tag{1.5}$$

1.8 Approaches and Contributions of the Thesis

The contributions of this thesis are summarized in the following subsections.

1.8.1. A Multiclass EEG Signal Classification Model Using Spatial Feature Extraction and eXtream Gradient Boosting Algorithm

Our first proposed optimization mainly focuses on mitigating the noise and unwanted peaks in the neural signals. In this experiment, we demonstrate the suitability of different data smoothing schemes on multiclass EEG signals. In the proposed study, we implement Savitzky - Golay curve smoothing approach to brain signals and computed the change in the noise-dispersion level before and after optimization. Thereafter, we normalized the improved signals and discriminated using the eXtreme Gradient Boosting (XgBoost) classifier. The computed classification accuracy was better than the other competitive methods and results of published BCI competition winners.

Part of this contribution has been published in:

Tiwari, A., & Chaturvedi, A. (2019, November). A multiclass EEG signal classification model using spatial feature extraction and XGBoost algorithm. In *2019* ***IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*** (pp. 4169-4175). IEEE.

1.8.2. Filter-based Channel Selection Approach Using Dynamic Channel Relevance

Our second experiment proposes a novel mutual information-based dynamic ranking procedure to sort the available EEG channels in the context of their relevance with candidate solutions. This scheme combines two information-theoretic paradigms: (1) Gini Index, and (2) Maximal Information Coefficient to determine the significance of an individual channel. This step effectively follows the “maximum relevance with minimum redundancy” principle while designing the optimal channel subset and maximizing the corresponding classification accuracy. Our experimental results show that the proposed channel ranking algorithm achieves better classification accuracy than the traditional approaches on all three BCI datasets. In addition, the proposed approach realizes a superior channel reduction rate by eliminating irrelevant and redundant channels which further helps to improve the practical applications of a BCI model.

Part of this contribution has been published in:

A. Tiwari and A. Chaturvedi, "A Novel Channel Selection Method for BCI Classification Using Dynamic Channel Relevance," in ***IEEE Access***, vol. 9, pp. 126698-126716, 2021, doi: 10.1109/ACCESS.2021.3110882.

1.8.3. Wrapper-based Channel Selection for Multiclass BCI Classification Using Multi-objective Improved Firefly Algorithm

In our third experiment, we introduce an improved Firefly Algorithm (FA), which can be applied not only for channel selection but other optimization problems. The proposed iterative methodology utilizes a bi-objective fitness function to investigate the quality of computed solutions. Because the proposed method explores a metaheuristic technique (FA) as an instance update strategy, it guarantees an optimal channel subset that improves the global performance of the BCI. The cross-validation results show that the proposed methodology obtains better classification accuracy and channel reduction rate as compared to various state-of-the-art models based on conventional metaheuristics and information-theoretic-based channel selection methods.

Part of this contribution has been published in:

Tiwari, A., Chaturvedi, A. Automatic EEG channel selection for multiclass brain-computer interface classification using multiobjective improved firefly algorithm. ***Multimedia Tools and Applications (2022) (Springer)***. https://doi.org/10.1007/s11042-022-12795-2.

1.8.4. Improvement in Wrapper-based Channel Selection Technique Using Multi-objective X-shaped Binary Butterfly Optimization Algorithm

In our fourth experiment, we extend our second experiment where the firefly algorithm was used to determine the optimal channel subset. In this methodology, we develop a novel binary variant of the conventional Butterfly Optimization Algorithm (BOA) using an X-shaped transfer function. It has already been proved that the BOA has a better solution search strategy than FA in terms of finding an optimal solution. This approach guarantees to improve the quality by applying uniform-crossover operation between the previous best solution and newly computed solutions.

The proposed methodology was validated on all three datasets and results were compared with different baseline channel selection techniques. The comparative study shows that the proposed methodology performs better than existing optimization schemes, especially in the case of large datasets.

Part of this contribution has been published in:

Tiwari, A., & Chaturvedi, A. (2022). Automatic Channel Selection using Multiobjective X-shaped Binary Butterfly algorithm for Motor Imagery Classification. ***Expert Systems with Applications (Elsevier)***, 117757.

1.8.5. A Hybrid Feature Selection Approach Based on Information Theory and Dynamic Butterfly Optimization Algorithm for Data Classification

In our fifth experiment, we develop a hybrid feature selection method by combining the Dynamic variant of the BOA (DBOA) and Feature Interaction Maximization (FIM) algorithms. The DBOA uses local search-based mutation operation to improve the standard of the produced solution. The DBOA was iteratively implemented on twenty datasets and FIM based feature selection approach was applied to select the best feature set. The quality of the computed solution was evaluated using a tri-objective fitness function (classification accuracy, relative number of selected features, and feature interaction score). This procedure provides the lower, average, and upper bound of the classification accuracy with the respective optimal feature set. The performance of the proposed feature selection method was compared with six conventional metaheuristic algorithms and four different variants of BOA. The performance of the proposed methodology was better on sixteen out of twenty datasets. In addition, the convergence rate, specificity, sensitivity, and feature reduction were also better than the state-of-the-art methods.

Part of this contribution has been published in:

Tiwari, A., & Chaturvedi, A. (2022). A hybrid feature selection approach based on information theory and dynamic butterfly optimization algorithm for data classification. Expert Systems with Applications (Elsevier), 196, 116621.

1.9. Thesis Organization

The remaining thesis is organized as follows. Chapter 2 presents the literature review and related works. The four research objectives discussed above are presented in the subsequent five chapters from Chapter 3 to Chapter 7. Chapter 8 concludes the thesis.

Chapter 3 presents a noise estimation strategy to reduce the unwanted peaks present in the neural signals and classify them using the boosting-based decision-tree model. Chapter 4 proposes a channel ranking procedure using a mutual information (MI) based criterion. In Chapter 5, an improved Firefly Algorithm (FA) based metaheuristic model was developed to reduce irrelevant and redundant channels. Chapter 6 proposes an X-shaped Binary Butterfly Optimization Algorithm (MX-BBOA) to improve the global performance of the BCI system. In Chapter 7, the hybrid-feature selection method was introduced to deal with the CoD problem. The thesis ends with Chapter 8 in which we sum up the obtained results and describe the work that can be done in the future to further improve the accuracy of the proposed algorithms.