Investigation of the Effects of Statistically Significant Features on the Classification of EEG-Based Motor Imagery Tasks

Submitted to the Graduate School of Natural and Applied Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Department of Biomedical Technologies

by

Murside Degirmenci

ORCID 0000-0003-0978-9653

November, 2023

This is to certify that we have read the thesis **Investigation of the Effects of Statistically Significant Features on the Classification of EEG-Based Motor Imagery Tasks** submitted by **Murside Degirmenci**, and it has been judged to be successful, in scope and in quality, at the defense exam and accepted by our jury as a DOCTORAL THESIS.

APPROVED BY:

Advisor: **Assoc. Prof. Dr. Yalcin Isler**
İzmir Kâtip Çelebi University

Co-advisor: **Assist. Prof. Dr. Yilmaz Kemal Yuce**
Alanya Alaaddin Keykubat University

Committee Members:

Prof. Dr. Savas Sahin
İzmir Kâtip Çelebi University

Prof. Dr. Mustafa Alper Selver
Dokuz Eylül University

Assoc. Prof. Dr. Yakup Kutlu
İskenderun Teknik University

Assist. Prof. Dr. Özlem Karabiber Cura
İzmir Kâtip Çelebi University

Date of Defense: November 20, 2023

Declaration of Authorship

I, **Murside Degirmenci**, declare that this thesis titled **Investigation of the Effects of Statistically Significant Features on the Classification of EEG-Based Motor Imagery Tasks** and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for the Doctoral degree at this university.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this university or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. This thesis is entirely my own work, with the exception of such quotations.
- I have acknowledged all major sources of assistance.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Date: 20.11.2023

Investigation of the Effects of Statistically Significant Features on the Classification of EEG-Based Motor Imagery Tasks

Abstract

Motor imagery (MI) task classification is highly prevalent in Electroencephalography (EEG)-based Brain-Computer Interface (BCI) research area. Extremity movement task classification and finger movement classification studies are presented in this thesis. In extremity movement classification, binary-class (right hand and left hand) and multi-class (right hand, left hand, right hand, and left hand) classifications are performed using 4 different feature extraction approaches and statistically significance-based feature selection (the independent t-test, one-way ANOVA test). Firstly, time-domain, Fourier Transform (FT)-based frequency-domain, and Wavelet Transform (WT)-based time-frequency domain features are calculated from multi-channel EEG signals. In addition to these features, Poincare plot measures-based non-linear features are calculated. Two different combination sets are also created to classify MI tasks of EEG segments using the extracted features. For finger movement classification, time-domain, frequency-domain, WT-based time-frequency domain, non-linear and their two different combinations set features are investigated using ANOVA-based and Pricipal Component Analysis (PCA)-based feature selection methods. Intrincsic Time-Scale Decomposition (ITD)-based time-frequency features are also investigated using ANOVA-based feature selection. 9 different machine learning algorithms namely Decision Tree (DT), Support Vector Machine (SVM), k-Nearest Neighbor (k-NN), Naive Bayes (NB), Logistic Regression (LR), Discriminant Analysis (DA), Neural Networks (NN), and Kernel Approximation (KA) are used based on 5-fold cross-validation to distinguish different groups. According to experimental results, the most successful feature sets are Poincare plot measures-based non-linear feature set and the combination set of different feature sets in extremity and finger movement classification studies. The statistically significance-based feature selection method improved classification performance in most of the feature sets.

Keywords: Extremity movement task classification, finger movement classification, intrinsic time-scale decomposition, motor imagery task, poincare plots, statistically significance-based feature selection

İstatiksel Anlamlı Özniteliklerin EEG Tabanlı Motor Hayali Görevlerin Sınıflandırmasındaki Etkisinin Araştırılması

Öz

Motor hayali (MH) görev sınıflandırması, Elektroensefalografi (EEG) tabanlı Beyin-Bilgisayar Arayüzü (BBA) araştırma alanında oldukça yaygındır. Bu tezde ekstremite hareketi görev sınıflandırması ve parmak hareketi sınıflandırma çalışmaları sunulmaktadır. Ekstremite hareketi sınıflandırmasında, ikili sınıf (sağ el ve sol el) ve çoklu sınıf (sağ el, sol el, sağ el ve sol el) sınıflandırmalar, 4 farklı öznitelik çıkarma yaklaşımı ve istatistiksel anlamlılığa dayalı özellik seçimi (bağımsız t-testi, tek yönlü ANOVA testi) kullanılarak gerçekleştirilmektedir. İlk olarak, Çok kanallı EEG sinyallerinden zaman alanı, Fourier Dönüşümü (FD) tabanlı frekans alanı ve Dalgacık Dönüşümü (DD) tabanlı zaman-frekans alanı özellikleri hesaplanır. Bu özniteliklere ek olarak Poincare çizimi ölçülerine dayalı doğrusal olmayan öznitelikler de hesaplanmaktadır. Çıkarılan öznitellikler kullanılarak EEG segmentlerinin MH görevlerini sınıflandırmak için iki farklı kombinasyon seti de oluşturulmuştur. Parmak hareketi sınıflandırması için zaman alanı, frekans alanı, WT tabanlı zaman-frekans alanı, doğrusal olmayan ve bunların iki farklı kombinasyon seti öznitelikleri, ANOVA tabanlı ve Temel Bileşen Analizi (TBA) tabanlı öznitelik seçim yöntemleri kullanılarak incelenmiştir. İçsel Zaman Ölçeği Ayrışımı (IZA) tabanlı zaman-frekans öznitelikleri, ANOVA tabanlı öznitelik seçimi kullanılarak da araştırılmaktadır. Karar Ağacı (KA), Destek Vektör Makinesi (DVM), k-En Yakın Komşu (k-EYK), Naive Bayes (NB), Lojistik Regresyon (LR), Ayırma Analizi (AA), Sinir Ağları (NN) ve Çekirdek Yaklaşımı (ÇY) farklı grupları ayırt etmek için 5-kat çapraz-doğrulamaya dayalı kullanılmaktadır. Deneysel sonuçlara göre ekstremite ve parmak hareketi sınıflandırma çalışmalarında en başarılı öznitelik setleri Poincare grafiği ölçümlerine dayalı doğrusal olmayan özellik seti ve farklı öznitelik setlerinin kombinasyon setidir. İstatistiksel anlamlılığa dayalı öznitelik seçme yöntemi, öznitelik setlerinin çoğunda sınıflandırma performansını iyileştirdi.

Anahtar Kelimeler: Ekstremite hareketi görev sınıflandırması, parmak hareketi sınıflandırması, içsel zaman ölçeği ayrıştırması, motor hayali görev, poincare çizimleri, istatiksel anlamlılığa dayalı öznitelik seçimi

To my family.

Acknowledgment

Murside Degirmenci was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) Scientist Support Department (BIDEB) under the 2211/A program number 1649B032001775 through the doctoral program.

This doctoral thesis was also supported by Izmir Katip Celebi University Scientific Research Council Agency as project number 2023-TDR-FEBE-0002 for Murside Degirmenci's doctoral thesis studies.

Murside Degirmenci was also supported by a research fellowship from the Higher Education Institution (YOK) within the 100/2000 Higher Education Institution Ph.D. scholarship.

Foreword

First of all, I would like to thank wholeheartedly my dear thesis advisor, Assoc. Prof. Dr. Yalcin Isler for his valuable support and encouragement throught my thesis work with his knowledge, patience, and understanding. I am lucky to have Assoc. Prof. Dr. Yalcin Isler as my advisor.

I also would like to thank my co-advisor Asst. Prof. Yilmaz Kemal Yuce for his valuable advice, support and quidance during his time as my co-advisor.

I would like to thank my thesis progress committee members, Prof. Dr. Savas Sahin and Assoc. Prof. Dr. Yakup Kutlu for their contributions to the progress of the thesis.

I also would like to thank my dear friend Lecturer Elif Izci for helping, advice, and valuable scientific insights.

Finally, my whole family, who went through this process with me, thank you from the deep of my heart. Thank you endlessly for your patience, constant support, love, and always being with me. I dedicate my dissertation to my family, my dear mother Nuray Değirmenci, my dear father Mehmet Erdal Değirmenci, and my dear twin Hakan Değirmenci. I'm so lucky to have you…

Mürşide Değirmenci

November, 2023

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACC	Accuracy
ADL	Activities of Daily Living
ANNs	Artificial Neural Networks
BiLSTM	Bidirectional Long Short-Term Memory
BOLD	Blood Oxygenation Level Dependent
BCI	Brain-Computer Interface
CNS	Central Nervous System
CSP	Common Spatial Pattern
CNNs	Convolutional Neural Networks
CV	Cross-Validation
DT	Decision Trees
DBS	Deep Brain Stimulation
DL	Deep Learning
DWT	Discrete Wavelet Transform
ECoC	Electrocorticography
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
EMD	Empirical Mode Decomposition
EL	Ensemble Learning
ERPs	Event Related Potentials
FN	False Negative

FP	False Positive
FFT	Fast Fourier Transform
FBCSP	Filter Bank Common Spatial Pattern
FIR	Finite Impulse Response
5F	Five Finger MI
FAWT	Flexible Analytic Wavelet Transform
FT	Fourier Transform
fMRI	Functional Magnetic Resonance Imaging
fNIRs	Functional Near-Infrared Imaging
GA	Genetic Algorithms
ITD	Intrinsic Time-Scale Decomposition
KA	Kernel Approximation
k-NN	k-Nearest Neighbours
LVQ	Learning Vector Quantization
LDA	Linear Discriminant Analysis
LR	Logistic Regression
LSTM	Long Short-Term Memory
MEG	Magnetoencephalography
MI	Motor Imagery
MIBIF	Mutual Information-based Best Individual Feature
MIRSR	Mutual Information-based Rough Set Reduction
NB	Naive Bayes
NBPW	Naive Bayesian Parzen Window
NNs	Neural Networks
NoMT	No MI Task Condition
CSP-OVR	One versus Rest CSP
ANOVA	One-Way Analysis of Variance

PSD	Power Spectral Density
PCA	Principal Component Analysis
PRCs	Proper Rotation Components
QDA	Quadratic DA
REM	Rapid Eye Movement
RCSP	Regularized CSP
STFT	Short-Time Fourier Transform
SCP	Slow Cortical Potentials
SSVEP	Steady-State Visual Evoked Potential
SVM	Support Vector Machine
TES	Transcranial Electrical Stimulation
tFUS	Transcranial Focused Ultrasound
TMS	Transcranial Magnetic Stimulation
TN	True Negative
TP	True Positive
2D	Two-Dimensional
VEPs	Visual Evoked Potentials
WPD	Wavelet Packet Decomposition
WT	Wavelet Transform

List of Symbols

Hz	Hertz
sec	Second
μV	Micro volt
μ	Mean value
SD	Standard deviation
δ	Delta
θ	Theta
α	Alpha
β	Beta
γ	Gamma
Ax	Activity
Mx	Mobility
Cx	Complexity
$Q1$	First quartile value
$Q2$	Inter-quartile value
$Q3$	Third quartile
$P(y(i))$	The probability that the signal is in the defined frequency domain
a_{ij}	Approximation coefficients of frequency band subsets
d_{ij}	Detail coefficients of frequency band subsets
L_t	The low-frequency component (baseline signal)
H_t	The high-frequency components (proper rotation components)
$\mathcal{L}$	The baseline extraction operator

$S(\omega k)$	Power spectral density
$X(\omega k)$	Discrete fourier transform of the signal
S_T	Total power
σ	Standard deviation
$\mathrm{d}(\mathrm{u}(\mathrm{i}), \mathrm{u}(\mathrm{j}))$	The distance between vectors u(i) and u(j)
λ	Eigenvalues
V	Eigenvectors
C	Covariance matrix
$\mathrm{P}(M_i)$	Class prior probabilities
P(X)	Prior probability of X sample
$P(X/M_i)$	The probability of X conditioned on M_i
$P(M_i/X)$	the probability of M_i conditioned on X
$\mathrm{P}_1(\mathrm{x})$	Probability value
β_0	The intercept

Chapter 1

Introduction

1.1 Background Information and Literature Review

1.1.1 Brain-Computer Interfaces (BCIs)

BCI, is a hardware and software communication system, ensures direct communication between the brain and a computer or external devices utilizing control signals obtained from signals of brain activity [1]. BCI systems use brain activity as input signals and then decode them to offer an extended degree of freedom improving life quality of persons who suffer from motor disabilities and serious physical disabilities [2]. BCI have been widely studies recently and have been different applications including rehabilitation, robotics, gaming, and neuroscience [3-5]. BCI applications have evolved substantially over the years and various publications have been introduced in the literature according to PubMed statistics as shown in Figure 1.1 [6].

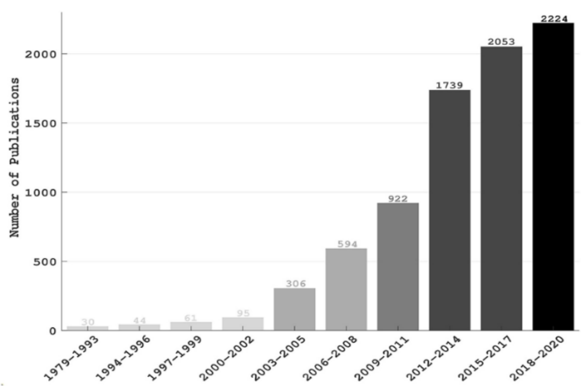


Figure 1.1: The number of publications over the years according to the PubMed statistics [6].

1.1.2 Brain-Computer Interfaces (BCIs) Applications

BCI system output could replace, restore, enhance, supplement, or improve natural CNS output to improve daily activities of paralyzed patients [2, 7-12]. In addition to these applications, a BCI system could affect interactions between the CNS and the external or internal environment of it. In another application of BCI, it could modulate brain signals using physical stimulation which are DBS, TES, TMS, tFUS, or other forms of brain signal modulation. These device-to-BCI connection applications are known as neuromodulation modalities [2]. The applications and main components of BCI system are given below in Figure 1.2 [2].

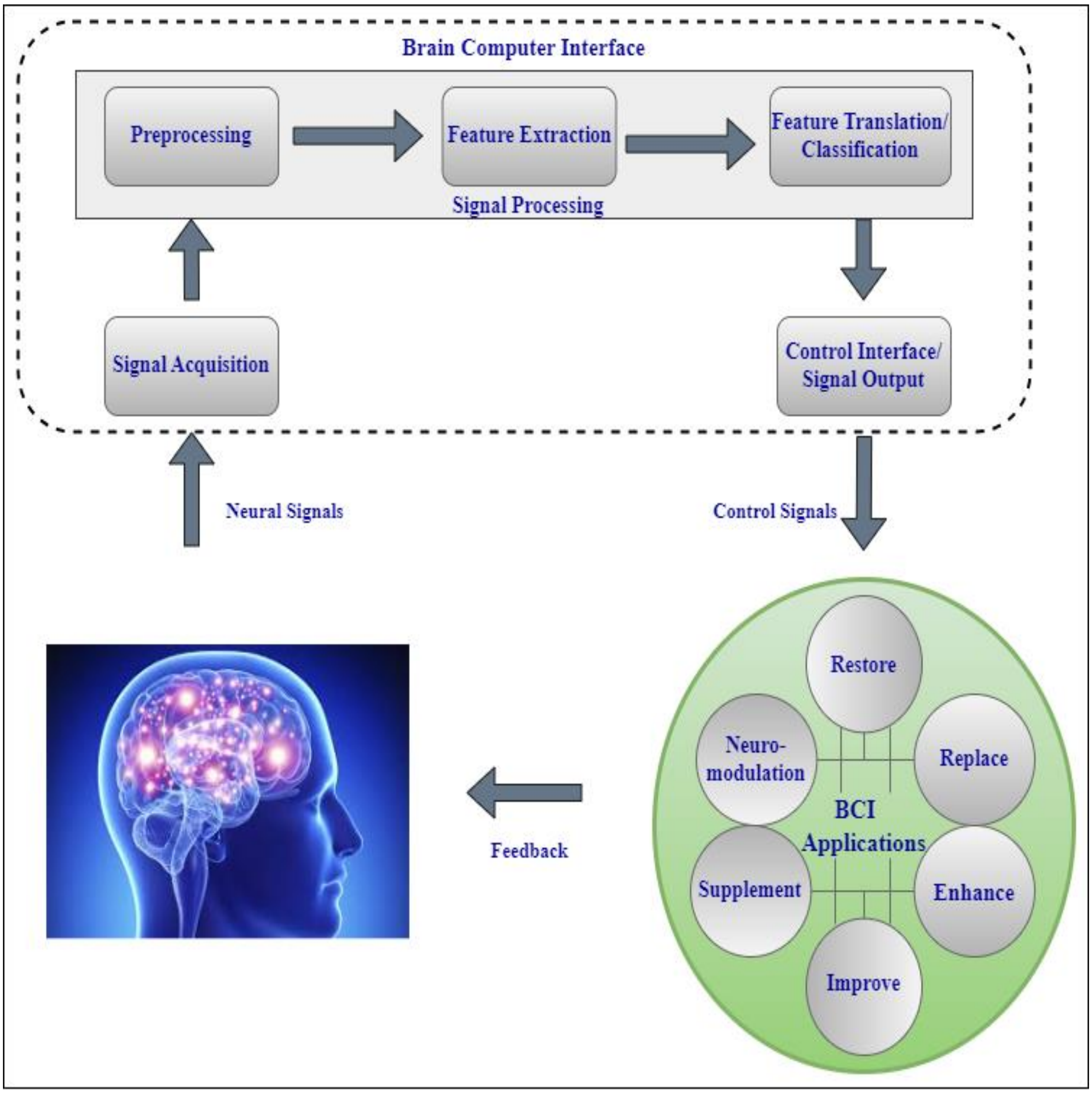


Figure 1.2: The main components and applications of BCI system.

1.1.3 Components of a Brain-Computer Interface (BCI) System

BCI systems are artificial systems that can recognize relevant and informative patterns in brain signals as the first step. After recording of brain signals, BCI systems extract features from signals and convert or translate the features into new outputs using signal processing step. Therefore, BCI systems consist of five main components such as signal acquisition, signal preprocessing, feature extraction, feature translation (classification), and control [2, 13]. Figure 1.2 shows the main components of BCI system.

The main components of a BCI system can be detailed as follows:

Signal acquisition: This stage, is the first step of BCI system, captures the activities (metabolic or electrophysiological) in brain and these recorded signals can be digitized for another signal processing stages. Various electrophysiological, and metabolic neuroimaging modalities can be used in brain signal acquisition to give an input for BCI system [1, 2]. Electrophysiological modalities are EEG, and ECoC, while metabolic modalities are fMRI, and fNIRs. The different advantages and disadvantages are available in each of them, however the electrophysiological modalities have been mostly preferred due to high temporal resolution and portability [1, 2].

Intra-cortical is an invasive technique and supplies high spatial resolution using electrodes which are implanted subdurally over brain cortex. It requires an operation for electrode placement [2, 14]. In another invasive brain imaging technique, which is known as ECoC, an effective representation of the underlying cortical electrical activity can be supplied with less invasiveness [15]. The non-invasive methods consist of EEG, MEG, fMRI, and fNIRs [2]. fMRI is a non-invasive brain imaging technique, and it measures BOLD response to capture brain activity. It provides relatively low temporal resolution and high spatial resolution [2, 16, 17]. fNIRs is another non-invasive technique that uses near-infrared light to measure blood flow dynamics to analyze neural activity. It has different advantages such as high spatial resolution, its portability and relatively cost-effective system. However, it provides low temporal resolution and less effective than based on electromagnetic signals [18, 19]. MEG is non-invasive imaging technique and measures the magnetic induction generated by electrical activity in neural cells. It requires specialized equipment and a laboratory environment. It is less cost-effective and more sophisticated than EEG-based BCI [2, 20, 21]. Among the different neuroimaging modalities, EEG is mostly used to acquire and feed input signals to BCI systems measuring the electrical activities of brain which results from the communication activity of neurons in the brain. It is mostly preferred for clinical and commercial use due to its high temporal resolution, non-invasiveness and inexpensive [2]. German psychiatrist Hans Berger is the first person who records the electrical field of human brain [22]. The electrodes are located on the surface of the scalp as stated in the international 10-20 electrode placement scheme as shown in Figure 1.3 [22, 23].

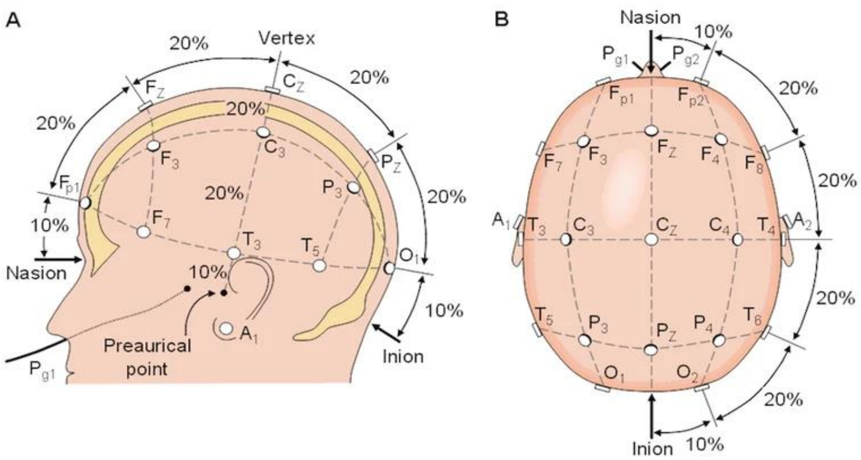


Figure 1.3: International 10-20 electrode placement system [22].

A standard EEG signal has the amplitude which ranges between 0-200µV, and the frequency of EEG signals differs between 0.5-50Hz. It consists of various frequency bands which are defined as delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ) from low frequency to high frequency, respectively [14]. The relevant characteristics, the frequency ranges, the waveforms of frequency bands, and their personal states vary amongst themselves. Delta waves have frequency ranges of 0.5-4Hz, and they are observed during deep sleep state in adults. The frequency ranges of theta waves are 4-8Hz, and these waves are observed in young and older children, sleeping adults, and sleep stages such as REM sleep. Alpha waves lie between frequency ranges of 8-13Hz, these waves are associated with awake, eyes closed, and resting states. The frequency ranges of beta waves are 13-30Hz, and they are related to motor activities. The personal states of beta waves are mental activations, and stress/anxiety situations. The gamma bands lie within the 30 to 100Hz range, the personal state of them is whole brain activity. Gamma waves are not preferred for EEG-based system design, because they are exposed to the artifacts which affect them [1, 14].

Signal preprocessing: This stage obtains a suitable form of signals for further signal processing of signals using different processes such as signal filtering, channel selection, and signal segmentation [2].

Feature extraction: In this stage, the discriminative and relevant patterns in the brain signals are captured using different features. The extraction of the effective features is a very important task in BCI systems. Since, the brain signals have non-stationary forms and these signals are subject to noise by artifacts such as EMG or EOG signals during signal acquisition. And also, the different feature selection methods can optionally be used to reduce feature dimension and complexity of system without the loss of relevant information. Therefore, effective feature extraction and feature selection methods play an important role in BCI design for using discriminative and effective features [2].

Feature translation (classification): The extracted features are classified using various machine learning algorithms to predict the corresponding brain activity and commands occurred to use in BCI [24].

Control: In this stage the classified brain signals are translated into relevant commands to control any connective device such as a wheelchair, a computer or a neuroprosthesis device [2,7-12, 25].

The various neural control signals are used by BCIs. These are sensorimotor rhythms, SCP, the P300 event-related potential, and ERPs such as VEPs and SSVEP [26, 27]. Sensorimotor rhythms have been analyzed in BCI research area. MI is traditionally established on visual or auditory feedback. MI task performed when subject only imagines moving any limb without actually moving any limb [28, 29]. Then sensorimotor rhythms are extracted and classified using different signal processing algorithms. Finally, the visual or auditory feedback is generated to the subject in regard to the success of system. The general concept of MI task in EEG-based BCI studies is represented as in Figure 1.4 [30].

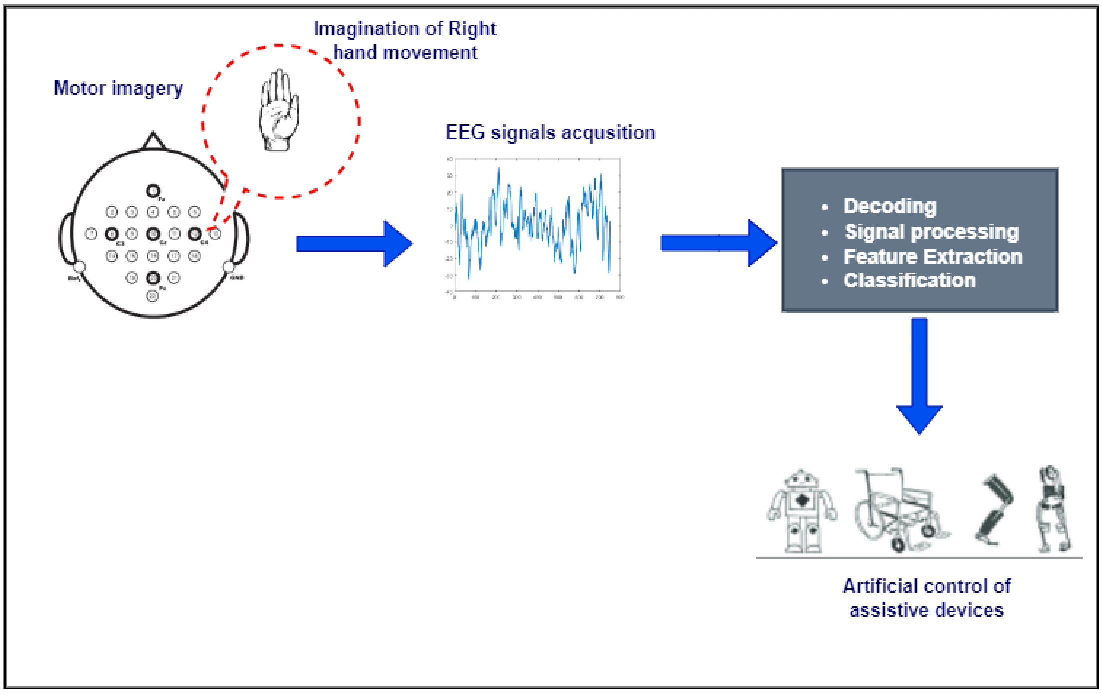


Figure 1.4: Motor imagery task in EEG-based BCI studies.

1.1.4 Extremity Movement Task Classification and Literature Review

In design of MI EEG-based systems, the extremity movement task classification studies have been generally performed rather than finger movement classification [31-33]. The extremity movement task generated when a subject only imagines the movement of different large limbs such as right hand, left hand, right feet, left feet, both feet, and tongue without actually moving any limb [33]. The accurate classification of extremity movement task is important to enable effective communication link which assists people suffering from motor activities by reason of any accident or disease [31, 34-35]. The automatic MI EEG signals classification has drawn attention in BCI studies and different signal processing approaches have been performed to sort out in research area of extremity movement task classification [36-38].

The main steps of MI EEG-based BCI systems are feature extraction, feature selection, and classification. Various studies have been carried out for extremity movement task classification using temporal, spectral, and spatial domain features. The statistical-based and amplitude-based time domain features, raw EEG time-series have been evaluated as temporal features [39-40]. Spectral features consist of frequency-domain [41], and time-frequency domain features [36], which are the mostly evaluated features for the analysis of MI EEG signals. FT represents the frequency domain of EEG time series and is one of the mostly used methods to obtain frequency distribution of EEG signals for extremity movement task classification. In 2017, authors [41] performed binary-class extremity movement classification study using FFT and LVQ networks. However, FT has some disadvantages, which it rules out non-stationary structure of EEG signals and does not include any time information in its frequency distribution [42]. Therefore, various time-frequency representations have been used to analyze MI EEG signals [43].

Various signal decomposition algorithms such as WT [43] and its derivatives [44, 45], STFT [36], ITD [46], EMD [43, 47], and its derivatives have been successfully used in discrimination of extremity movement tasks. Ha and Jeong [36], presented STFT-based approach for binary-class extremity movement classification. EEG signals were converted into 2D images and these EEG representations applied in CNNs architecture for classification. In 2018, Alam and Samanta [47] performed EMD-based MI task classification study. In [45], authors presented MI task classification study utilizing DWT and cross-correlated EEG features. In a different approach [44], Chaudhary et al. performed binary-class (right hand and right foot) MI task classification study using the FAWT approach. They decomposed the MI EEG signals into sub-bands and temporal-moment based features were obtained from these bands. EL-based classification was performed with promising results. In another study [43], MI task classification study was performed using WPD, k-NN algorithm and higher-order statistical features. In 2018, Mohamed et al. [46] carried out four-class MI task classification study using ITD and ANNs algorithm. They extracted the PRCs using

ITD and evaluated the energy, entropy, and mean absolute values from PRCs as features.

After a brief investigation, it was observed that spatial features have been mostly evaluated to analyze MI EEG signal with promising classification results in last decades. CSP [48] and derivatives such as FBCSP [49] are the most applied methods to obtain spatial features for the extremity movement task classification. In [50], CSP and DL approaches were employed to distinguish MI EEG signals. In a different approach [51], CSP-based and wavelet coherence-based feature extraction processes were conducted for binary-class MI EEG signal classification. Lu et al. [52], used the aggregated RCSP for analysis of MI EEG signals. In another study [49], Ang et al. proposed FBCSP-based binary-class and multi-class extremity movement task classifications. They also applied feature selection using MIBIF and the MIRSR algorithms. NBPW algorithm is applied to classify MI EEG signals.

In addition to the effectiveness of feature extraction methods, the correlation of these features plays an important role in improving the classification performance. Additionally, the high amounts of features increase the complexity of classification process due to redundant information [53, 54]. In the literature, various feature selection methods have been introduced and used in EEG signal processing to obtain relevant and effective features by selecting of features. The backward elimination, PCA, GA, and statistically-significance based feature selection methods have been mostly applied in literature to improve classification performance selecting effective features and diminishing computational load of classifiers [53-55].

According to the studies, the most successful experimental results have been reported using machine learning algorithms such as SVM [40], ANN [46], k-NN [43], NB [49], and EL [40, 44] and various DL models [36, 50].

After a brief investigation of extremity movement task classification studies, some drawbacks and limitations of studies can be summarized as follows;

- The succesfully employed feature extraction methods were spectral and spatial feature extraction-based approaches for analysis of MI EEG signals. In addition to these features, the effectiveness of various time-domain, non-linear, time-frequency domain features and the different combinations which consist of variety of features can be investigated to improve the classification performance.

- It was observed that the feature selection algorithms have not been mostly applied to analyze MI EEG signals. Various feature selection methods such as statistically significance-based feature selection and PCA, which is successfully applied in EEG signal processing can be improved to the classifier performance.
- In addition to the successful machine learning such as SVM, ANN, k-NN, NB, and EL, the effectiveness of different algorithms such as DT, LDA, NNs, LR and KA can be analyzed for extremity movement task classification.

1.3.5 Finger Movement Task Classification and Literature Review

Rehabilition of motor functions of a hand, especially fingers is an important task to improve ADL for humans who exposure to upper limb motor impairment [56, 57]. Finger movements are essential tools to manipulate and move objects and interact with environment [58]. Hands consist of various types of tissues such as skeletal muscles, bones and joints [59, 60]. The sophisticated finger movements are required complex processing in central nervous system [56, 61]. Especially, the motor activities of hand can be critically affected after stroke, which affects approximately 100-200 out of every 100,000 people, and is the major cause of motor disability [62, 63]. Furthermore, these deficiencies in ADL considerably affect the patient' independence and also cause long term disability [56]. Therefore, the accurate decoding of finger movement is an important task and can help people who suffer from motor disabilities by improving ADL. In recent years, finger movement classification has become a very important research topic and various signal processing algorithms have been used to solve this task [56-57, 64].

In literature, EEG-based finger movement classification studies have been introduced utilizing various signal processing and classification methods. Different types of features including temporal, spectral and spatial features are utilized to improve the classification performance of EEG segments. Various methods have been introduced for classification of finger movement by using temporal features such as the raw EEG time series [65-67], different amplitude-based and statistical features. As spectral features, FT is one of the most exploited methods for EEG analysis and it has been frequently applied to this task [66]. Additionally, various time-frequency representations of EEG signals including WT [68], STFT [69], EMD [70] and its derivatives have been successfully utilized in classification. The spatial features such as CSP [71, 72] and its derivatives [73] have been mostly used to classify finger movement and successful classification results have been supplied for finger movement classification.

Kaya et al. [66] proposed a five-finger movement classification study utilizing time-domain and frequency domain features such as power of EEG sub-bands, FT amplitudes and EEG time series. They performed both subject-dependent and subject-independent analysis with 19 channel EEG signals of 8 subjects. SVM classification algorithm is applied for classification. In another study [71], five finger movements are classified using spatial features which are extracted with CSP algorithm. The subject-dependent analysis is performed with 4 subjects. The extracted features are classified with RF algorithm. In 2022, Azizah et al. [69] conducted a channel process-based analysis using CSP-OVR and 4 out of 19 EEG channels are selected before feature extraction process. They evaluated spectral features obtaining spectrogram features from the chosen EEG channels. The subject-dependent analysis is carried out using SVM algorithm with promising classification results. In [73], authors performed subject-dependent finger movement classification using spatial (multi-class CSP) and spectral features (complex Fourier amplitudes). These features are evaluated from 19 EEG channels and classified with SVM algorithm.

In addition to traditional machine learning algorithms, deep learning approaches have also been used to improve classifier performance in finger movement research area. Various deep learning approaches such as CNNs [65, 67, 74], LSTM and their different variants [70] are utilized to discriminate finger movements. Mwato-Velu et al. [70] presented an EMD-based subject-dependent classification study utilizing 4 selected and relevant channels of 19-channeled EEG signals. A deep learning model which is known as BiLSTM is adopted to classify finger movement. In 2022, authors [67] presented a deep learning-based subject-dependent finger movement classification study using EEG time series. Before feature extraction, EEG signals of selected 4 subjects are used and 4 out of 19 EEG channels are selected for analysis. In study [72], CSP algorithm-based feature extraction and deep learning approach are used to perform a subject-dependent classification. The experimental analysis performed using 19-channel EEG signals of 4 subjects. In a recent study [65], CNN-based classification approach is introduced for subject-independent classification. As feature extraction process before giving to CNN structure as input data, EEG time series combined with sliding window and noise enhancement methods. The 19-channel EEG signals of 8 subjects are utilized for signal processing. In another recent study performed by Limbaga et al. [74], a CNN architecture is employed to EEG signals both feature extraction and classification. In addition to their proposed study, a transfer learning model is applied to reinforce their deep learning model. They performed subject-independent analysis using 14 channels out of 19-channel EEG signals of 4 subjects.

According to literature studies, in classification stage, the successful classification results have been reported by utilizing different machine learning algorithms including SVM [66, 69, 73, 75], RF [71] and EL [75], etc. In addition to these machine learning algorithms deep learning approaches such as CNN [65, 67, 74] and LSTM [70] architectures have been successfully used for finger movement classification.

Considering the recent literature studies performed for finger movement classification, the main benefits, drawbacks and difficulties of this task can be listed as follows;

- The classification results remained at low rates when all EEG channels were used for analysis and the subject-independent analysis were performed. According to studies, channel selection and subject-dependent analysis can be used to improve classification performance.
- After a brief comparison of extremity classification and finger movement classification studies, it was observed that the classification results of finger movement classification studies have remained at low rates than the classification results of extremity classification studies due to complex neural processing. The selection of the effective feature extraction and classification methods can play an important role in improving classification results, especially in finger movement classification.

- Many studies have frequently used time-domain, frequency-domain and spatial-domain features with promising classification results. The effectiveness of non-linear parameters, various time-frequency algorithms, and different combinations of features varieties in finger movement classification can be investigated to improve classification performance.
- In the studies carried out for finger movement classification, the feature selection methods have not been generally included in the processing of EEG signals. The effective feature selection methods can be improved classifier performance defining relevant and fewer features and reducing the classifier complexity.
- According to literature studies, it was observed that one of the most successful classification algorithms is SVM. In addition to SVM, various classification algorithms such as DT, LDA, NB, k-NN, EL, NNs and KA can be investigated to improve classification performance and their effectiveness can be analyzed and compared.

Therefore, the effectiveness of various types of features and their combinations can be investigated with effective feature selection methods and various classifier algorithms to improve classification performance for finger movement discrimination.

1.2 Objectives of the Thesis

The main purpose of the presented thesis is to obtain high classification results with various signal processing methods for two MI EEG signal classification, extremity movement task classification, and finger movement classification, which are frequently analyzed in the literature BCI studies. The various EEG signal processing methods are used for classification of extremity movement task classification, and finger movement classification. The EEG segments of MI tasks were investigated using various feature sets such as time-domain feature set (amplitude-based and statistical based features), frequency-domain feature set (FFT-based features), time-frequency domain features (WT-based features) and non-linear feature set (Poincare features) to improve classification performance. In addition to these feature sets, time-frequency features extracted using ITD algorithm were investigated for finger movement classification.

The main objectives of the thesis can be given as follows:

1. Distinguishing of MI EEG signals using the various feature sets separately and their different combinations was performed to analyze the two MI EEG signal classification. The effectiveness of all feature sets was investigated for both binary-class and multi-class MI task classification in extremity movement task classification.
2. In finger movement classification, experimental analysis was carried out for both subject-dependent and subject dependent analysis.
3. We investigated whether the recently presented ITD can be applied for representation and classification of multi-channel MI EEG signals.
4. We aimed to investigate the effectiveness of statistically-significance based feature selectin method to improve classification model.

1.3 Contributions of the Thesis to Literature

This thesis aimed to present various signal processing methods from the literature for both extremity movement task classification and finger movement classification using different features, effective feature selection methods and various machine learning algorithms.

The main and innovative contributions of the studies of this thesis can be highlighted as follows:

a) For extremity movement task classification approaches;

1. We investigate the effectiveness of hand-crafted feature extraction methods considering various feature sets such as time-domain, frequency-domain, time-frequency domain, non-linear feature sets and different combinations of these feature sets.
2. For the first time, novel non-linear features from the Poincare plot measures of MI EEG signals are implemented in this study. Thus, we show that non-linear feature extraction-based approach provides promising experimental results and it is an efficient method for classification of MI EEG signals.
3. We demonstrated the effectiveness of the statistically-significance based feature selection methods by comparing the experimental results of analyzes

using all features and statistically significant features determined by feature selection.

4. The effectiveness of all feature sets and the statistically-significance based feature selection methods are investigated for both binary-class and multi-class extremity movement task classification.
5. This thesis is the first study that performed analyzes and comparison of various machine learning algorithms in extremity movement task classification to the author’s best knowledge.

b) For finger movement classification approaches;

1. The various features including time-domain, frequency-domain, time-frequency domain, non-linear features and their different combinations are used for analysis of EEG signals of finger movement.
2. For the first time, Poincare plot-based non-linear features are extracted for finger movement classification in addition to the traditional features.
3. In addition to different features which are also evaluated in extremity movement task classification, a different approach is conducted using ITD-based features for only finger movement classification. The first three PRCs are extracted from EEG signals, the effectiveness of these components and their different combinations are investigated using different features, separately. To the author's best knowledge, this is the first study that investigated the effectiveness of different PRCs and their combinations for finger movement classification. Here we demonstrate that the ITD-based approach can be successfully utilized to analyze MI EEG signals and the proposed method, combination of PRCs, improved the classifier performance.
4. It has been noted that the statistically-significance based feature selection provides successful classification of finger movement in analyzes using various feature sets.
5. In contrast to the traditional finger movement classification studies which have ignored passive mode (NoMT), we performed a six-class finger movement classification study implementing EEG signals NoMT task to design a more realistic BCI design for patients, who suffering from motor disabilities to the author's best knowledge.

6. The subject-dependent and subject-independent analyzes are carried out using different features and machine learning algorithms.
7. This thesis is the first study that performed analyzes and comparison of various classifiers in finger movement classification to the author's best knowledge.

Chapter 2

Materials and Methods

In this chapter, experimental data sets, various feature extraction methods, and feature selection methods used for signal analysis, and machine learning algorithms applied for classification section of proposed feature sets are listed. These are introduced in the following sections.

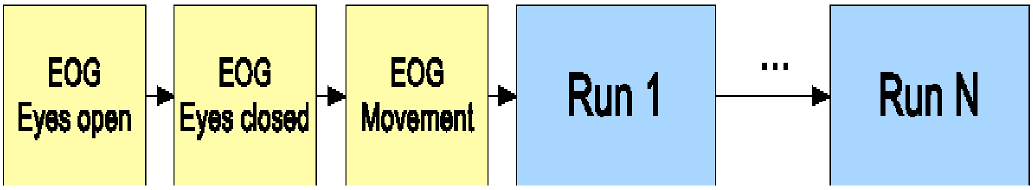
2.1 Experimental Data Sets

In this thesis, two different EEG data sets are used to perform experimental section. The first one is BCI Competition IV Dataset IIa that is publicly available MI EEG dataset generated during MI tasks. These MI EEG signals are utilized for extremity movement classification analyzes in our thesis. The second one is a publicly available EEG dataset that is a large electroencephalographic MI dataset for EEG-based BCIs. Multi-channel EEG signals of finger movements from this dataset are evaluated for finger movement classifications in our thesis.

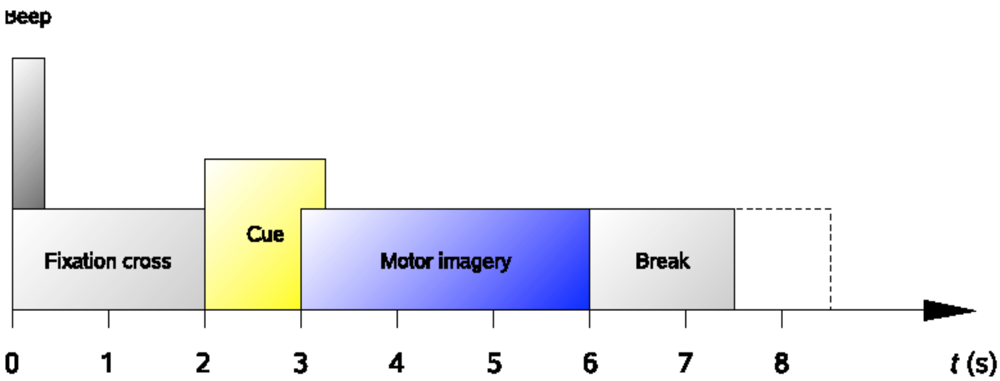
2.1.1 BCI Competition IV Dataset II-a

In this thesis, the effectiveness of proposed approaches for extremity movement classification is evaluated using BCI Competition IV Dataset IIa [76]. It includes 22-channel EEG data collected form 9 subjects (4 female and 5 male). The cue-based BCI paradigm includes different MI tasks the imagination of the left hand (class 1), right hand (class 2), both feet (class 3), and tongue (class 4).

EEG data was recorded as two sessions on different days. Each session includes 6 runs which are divided by breaks. And also, each run includes 48 trials which are categorized 12 trials for each class category. A total of 288 trials were recorded at the end of all runs for each subject. In experiments, a cue which is one of the four MI classes is represented to the subject during 1.25 sec. Thus, the subject is directed to perform the requested task. Subject performed related MI task until a shorth break. Therefore, the MI task is performed during 3 sec for each trial. format throughout the manuscript. The timing schemes of experiment are represented in Figure 2.1 [76].



(a) Session timing representation



(b) Trials representation.

Figure 2.1: The timing scheme for (a) One session, (b) The Paradigm [76].

In data recording, the EEG signals were sampled with 250 Hz and filtered with a band-pass filter between 0.5 Hz and 100 Hz. In addition, the 50 Hz notch filter was applied to extinguish line noise. As preprocessing section of extremity movement classification studies, MI EEG segments of EEG signals for each trial are divided for signal processing.

2.1.2 A large electroencephalographic motor imagery dataset for electroencephalographic brain computer interfaces

In this thesis, the effectiveness of proposed approaches for finger movement classifications is evaluated using a large electroencephalographic MI dataset for EEG-based BCIs which is introduced by Kaya et al. [66]. The data set comprised of various type of MIs in 4 different paradigms. 21-channel EEG signals were recorded from 13 healthy subjects using EEG-1200 JE-912A recording system. According to the international 10-20 EEG electrode placement system, 19 EEG, 2 reference electrodes and the ground electrode were located for experiments. BCI interaction paradigms are designed based on MI tasks of 10 different limbs in this data set are listed as follows:

- Paradigm #1 (CLA): It is defined as classical left/right hand MI model and includes three MIs of left and right-hand movements and one passive mental imagery that is defined as passive model and any MI task is not available in it.
- Paradigm #2 (HaLT): It is an extension form of 3-state CLA paradigm model and it includes the imagery of left and right leg movements and tongue movement. Therefore, a total of six different MIs are available in this model.
- Paradigm #3 (5F): It is introduced as 5 finger MI and it consists of finger movements imageries which are MI EEG signals recorded during imagination of the movements of the five fingers on a hand as flexion of the corresponding fingers up or down. Finger movements imageries are denoted as Thumb (Class 1), Index finger (Class 2), Middle finger (Class 3), Ring finger (Class 4), and Pinkie finger (Class 5).
- Paradigm #4 (NoMT): It is defined as no MI model or visual signals. In this paradigm no visual stimulus is shown on the screen for subjects, they passively watched the screen.

In our finger movement classification analysis, six different class categories are available utilizing 5F and NoMT paradigms. In the recording of EEG signals, the action signal is represented to subjects during 1 sec, subject implemented desired motor imagery in this time. Then, the related action signal is not remained on the screen and 1.5-2.5 seconds break is given for subjects until the next motor imagery action signal is given. Two different sampling frequencies including 200 Hz and 1000 Hz are available for EEG signals in this dataset. In this thesis, EEG signals obtained with a 1000 Hz sampling frequency were selected to be utilized for signal analysis of finger movement classification. In the acquisition of EEG signals, a 0.53 Hz-100 Hz band-pass filter was implemented to the signals utilizing hardware filters. In order to suppress the electrical grid interface, a 50 Hz notch filter was used to signals.

In the preprocessing section of finger movement classification of this thesis, 100 samples of 1000 Hz EEG signals of six different classes MIs (5F and NoMT paradigms) were determined to be used in signal processing and following classification section. The same number of EEG segments were determined for each of six classes to provide balanced data distribution for analysis. Therefore, 600 trials are available for each subject.

2.2 Feature Extraction

In this thesis, six different feature sets are utilized for the classification of extremity movement task and finger movement. MI EEG signals are investigated utilizing (i) time-domain feature set, (ii) frequency-domain feature set, (iii) time-frequency domain feature set, (iv) non-linear feature set, (v) combination of time-, frequency-domain feature sets and time-frequency domain feature set, and finally, (vi) combination of time- and frequency-domain feature sets, time-frequency domain feature set and non-linear feature set to provide high classification performance.

2.2.1 Time Domain Feature Set

In this thesis, 24 different time-domain features are extracted using original time-domain information of EEG signals. These features are evaluated according to the amplitude and statistical changes of the EEG signal. 24 different time-domain features and their mathematical formulas are listed as follows [40, 75, 77-79]:

$$Minumum\ value\ =\ \min\left(X[n]\right) \tag{2.1}$$

$$Maximum\ value\ =\ \max\left(X[n]\right) \tag{2.2}$$

$$Mean\ (\mu)\ =\ \frac{1}{N}\sum_{n=0}^{N-1} X[n] \tag{2.3}$$

$$S\tan d\,art\ deviation\ value\ =\ \sqrt{\frac{1}{N}\sum_{n=0}^{N-1}(X[n]-\mu)^2} \tag{2.4}$$

$$Integrated\ EEG\ value\ =\ \sum_{n=0}^{N-1}|X[n]| \tag{2.5}$$

$$Mean\ absolute\ value\ =\ \frac{1}{N}\sum_{n=0}^{N-1}|X[n]| \tag{2.6}$$

$$Simple\ square\ integral\ =\ \sum_{n=0}^{N-1}|X[n]|^2 \tag{2.7}$$

$$Variance\ =\ \frac{1}{N-1}\sum_{n=0}^{N-1}(X[n]-\mu)^2 \tag{2.8}$$

$$Root\ mean\ square\ =\ \sqrt{\frac{1}{N}\sum_{n=0}^{N-1}X[n]^2} \tag{2.9}$$

$$Waveform\ length\ =\ \sum_{n=1}^{N}|X[n]-X[n-1]| \tag{2.10}$$

$$Average\ amplitude\ change\ value\ =\ \frac{1}{N}\sum_{n=1}^{N}|X[n]-X[n-1]| \tag{2.11}$$

$$Absolute\ difference\ in\ SD\ =\ \sqrt{\frac{1}{N}\sum_{n=1}^{N}(X[n]-X[n-1])^2} \tag{2.12}$$

$$Kurtosis\ =\ \frac{\frac{1}{N}\sum_{n=0}^{N-1}(X[n]-\mu)^2}{\left(\frac{1}{N}\sum_{n=0}^{N-1}(X[n]-\mu)^2\right)^2} \tag{2.13}$$

$$Skewness\ =\ \frac{\frac{1}{N}\sum_{n=0}^{N-1}(X[n]-\mu)^3}{\left(\sqrt{\frac{1}{N}\sum_{n=0}^{N-1}(X[n]-\mu)^2}\right)^3} \tag{2.14}$$

$$Hjorth\ parameters\ (Activity)\ =\ A\ =\ \frac{1}{N-1}\sum_{n=0}^{N-1}(X[N]-\mu)^2 \tag{2.15}$$

$$Hjorth\ parameters\ (Mobility)\ =\ M\ =\ \sqrt{\frac{A\left(\frac{dX\,[n]}{dn}\right)}{A\,(X[n])}} \tag{2.16}$$

$$Hjorth\ parameters\ (Complexity)\ =\ C\ =\ \frac{M\left(\frac{dX\,[n]}{dn}\right)}{M\,(X[n])} \tag{2.17}$$

$$Signal\ range\ =\ max\,(X[n])-\min\,(X[n]) \tag{2.18}$$

$$First\ quartile\ value\ (Q1)\ =\ X\left[\frac{N+1}{4}\right] \tag{2.19}$$

$$Inter - quartile\ value\ (Q2) = Q3 - Q1 = X\left[\frac{3(N+1)}{4}\right] - X\left[\frac{(N+1)}{4}\right] \quad (2.20)$$

$$Third\ quartile\ (Q3) = X\left[\frac{3(N+1)}{4}\right] \quad (2.21)$$

$$Mode\ value = Most\ frequent\ value\ in\ an\ EEG\ signal. \quad (2.22)$$

$$Slope\ change\ value = \sum_{n=1}^{N-1}\left[f[(X[n]-X[n+1]).(X[n]-X[n+1])]\right] \quad (2.23)$$

$$f(x) = \{1,\ \ if \rightarrow x \geq threshold,\}\ \{0,\ otherwise.\}$$

$$Zero - cros\sin g\ \ value = \frac{1}{2N}\sum_{n=1}^{N-1}|sign\,[X_i(n)] - sign\,[X_i(n-1)]| \quad (2.24)$$

$$sign\big(X_i(n)\big) = \{1,\ X_i[n] \geq 0\,,\}\ \{-1,\ X_i[n] < 0\,.\}$$

In the above equations, X[n] denotes the EEG signal and N denotes the size of the EEG signal. Mean value is denoted as μ.

2.2.2 Frequency Domain Feature Set

To provide this feature set, the frequency distribution embedded in the EEG signal is generated using FT. The various sub-bands of EEG embedded in signal which are delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ) are extracted from the EEG signals. Frequency ranges of these bands are defined as δ (0.5-4 Hz), θ (4-8 Hz), α (8-13 Hz), β (13-30 Hz), and γ (30-100 Hz). The relevant and discriminative frequency domain features such as energy, variance, and entropy are evaluated from the defined EEG sub-bands. The change of energy, variance, and entropy (irregularity) values in the defined frequency bands of EEG signal can be analyzed using these features. Calculation of energy, variance, and entropy are given in the following equations; respectively [78-81]

$$Energy_f = \sum_{i=1}^{M} y(i)^2 \quad (2.25)$$

$$Variance_f = \frac{1}{M-1}\sum_{i=1}^{M}(y_k - \overline{y})^2 \quad (2.26)$$

$$Entropy_f = \frac{1}{log\,(M)}\sum_{i=1}^{M} P\big(y(i)\big)\log\Big(P\big(y(i)\big)\Big) \quad (2.27)$$

where, f indicates the frequency band type (δ, θ, α, β, and γ) of EEG signals. The energy of these bands is evaluated using the power spectrum of signals. y indicates the FT of a real discrete time EEG signal and M indicates the maximum frequency. $\overline{y}$indicates the average of the y signal. Entropy is defined as irregularity and measures the regularity of the power spectrum of the EEG signal. P(y(i)) denotes the probability that the signal is in the defined frequency domain. EEG signal.

2.2.3 Time-Frequency Domain Feature Sets

We applied WT- and ITD-based approaches to analyze EEG signals in our study. WT-based time-frequency set is investigated for both extremity movement task classification and finger movement classification. ITD-based time-frequency feature set is investigated for only finger movement classification in this thesis.

2.2.3.1 Wavelet Transform-based Time-Frequency Domain Feature Set

In this thesis, the time-frequency domain features are evaluated using WT. It is mostly used analysis of non-stationary EEG signals by preserving the time-frequency resolution [82]. It is a smooth and fast oscillating function that is well localized in frequency and time. It can be used as particularly generated FIR filter [83-85]. The input EEG signal X[n] is decomposed into sub-frequency components [82-86]. The high frequency and low frequency components of the EEG signals are generated based on the frequency responses of the FIR. In each of the decomposition levels, both high-pass and low-pass filters are applied to the signal. Among these filters, the high-pass filter is related with the mother wavelet function [79]. The output of high-pass filters is known as detail (d) coefficients signals, while the output of low-pass filters is known as approximation (a) coefficients [82, 87]. The approximation signal is re-decomposed until the decomposition level is completed. In analysis of extremity movement, EEG signals with 250 Hz sampling frequency are analyzed using 7 level decomposition. In analysis of finger movement, EEG signals with 1000 Hz sampling frequency are analyzed using 9 level decomposition.

In WT-based analysis, one of the important steps is the defining of the mother wavelet function [84]. Various mother wavelet types are analyzed for EEG signal analysis in the literature. These are haar, db, sym, coif, bior, rbio, meyr, mexh, morl, cmor, and dmey [79, 83-85]. In this thesis, wavelet function is defined as Haar for both extremity movement task and finger movement classification. The frequency sub-bands (delta, theta, alpha, beta, gamma) are decomposed from MI EEG signals utilizing Haar mother wavelet. Then, the energy, variance, and entropy values of these sub-bands are evaluated as time-frequency features [79, 86, 88-92]. The energy at each decomposition level is evaluated according to the below equations:

$$Energy_{d_i} = \sum_{j=1}^{N} \left|d_{ij}\right|^2, \quad i = 1, 2, 3, \dots, l \tag{2.28}$$

$$Energy_{a_i} = \sum_{j=1}^{N} \left|a_{ij}\right|^2, \quad i = 1, 2, 3, \dots, l \quad (2.29$$

Here, (d_i) and (a_i) are utilized to provide subsets of EEG frequency bands (δ, θ, α, β, and γ) based on the decomposition tree. a_{ij} and d_{ij} indicate the (a) and (d) of frequency band subsets, respectively. The decomposition level is represented with l and i changes from 1 to l for calculations. The entropy values of each decomposition level are evaluated according to following mathematical formula:

$$Entropy_i = -\sum_{j=1}^{N} d_{ij}^2 \log\left(d_{ij}^2\right), \quad i = 1, 2, 3, \dots, l \tag{2.30}$$

In another feature variance value is evaluated utilizing following equation:

$$Variance_i = \frac{1}{N-1} \sum_{j=1}^{N} \left(d_{ij} - \mu_i\right)^2, \quad i = 1, 2, 3, \dots, l \tag{2.31}$$

$$\mu_i = \frac{1}{N} \sum_{j=1}^{N} d_{ij}, \quad i = 1, 2, 3, \dots, l \tag{2.32}$$

where, the mean value of decomposition level is indicated as $\mu_i \dots$

2.2.3.2 Intrinsic Time-Scale Decomposition-based Time-Frequency Domain Feature Set

ITD is the iterative signal decomposition algorithm and is introduced for the analysis of non-stationary and non-linear biomedical signals [93]. It divides the original signal into a sum of PRCs and a monotonic trend without using laborious and unproductive sifting or splines. The original signal is decomposed into low-frequency component (baseline signal) which is indicated as L_t and high-frequency components (proper rotation components) which are indicated as H_t [93-95]. Firstly, an EEG signal which is indicated as X_t defined for ITD analysis. $\mathfrak{L}$ is an operator, is determined to decompose baseline signal from original signal X_t and leave behind high-frequency components PRCs. Therefore, X_t can be represented as follows:

$$X_t = \mathcal{L}X_t + (1 - \mathcal{L})X_t = L_t + H_t \tag{2.33}$$

Here, the baseline signal is defined as $\mathcal{L}X_t$ and PRC is denoted as $H_t = (1 - \mathcal{L})X_t$. The process of the ITD algorithm to extract baseline and PRCs can be applied as given in **Algorithm** [93-95].

Algorithm: ITD

- A signal X_t which is available $t \geq 0$ its local extremes $\tau_k,\ k = 1,\ 2,\ \ldots$ is assumed for analysis. $X(\tau_k)\ \equiv X_k$ and $L(\tau_k)\ \equiv L_k$ notations are defined.
- The L_t and H_t are generated over the interval $[0,\ \tau_k]$, and the signal X_t is existed on $[0,\ \tau_k + 2]$. The baseline extraction operator $\mathcal{L}$ is introduced as piece-wise linear function between two extreme locations on the defined interval $(\tau_k,\ \tau_k + 1]$ according to following equations:

$$L_t = L_k + \left(\frac{L_{k+1} - L_k}{L_{k+2} - L_k}\right)(X_t - X_k), \quad t \in (\tau_k,\ \tau_{k+1}] \tag{2.34}$$

where

$$L_{k+1} = \alpha\left[\left(X_k + \frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}\right)(X_{k+2} - X_k)\right] + (1 - \alpha)X_{k+1}, \tag{2.35}$$

and $0 < \alpha < 1$, α is typically defined as $\frac{1}{2}$. The monotonicity of X_t is preserved using this method of obtaining L_t.

- After the extraction of L_t, the residual or PRC is evaluated as:

$$\mathcal{H}X_t = (1 - \mathfrak{L})X_t = H_t = X_t - L_t \tag{2.36}$$

Therefore, the original X_t signal can be reconstructed utilizing the decomposed L_t and H_t modes as follows:

$$X_t = L_t^D + \textstyle\sum_{k=0}^{D} H_t^k \tag{2.37}$$

Here, D denotes the number of decomposed high-frequency components (PRCs).

Various features are calculated from these ITD-based decomposed high-frequency PRCs. In our thesis studies, the first three PRCs (PRC1, PRC2, and PRC3), their binary combinations (PRC1–PRC2, PRC1–PRC3, and PRC2–PRC3), and their triple combination (PRC1-to-3) are utilized to evaluate 10 different features. The effectiveness of obtained feature sets is investigated separately.

10 different features such as **mean, higher order moments (1st, 2nd, 3rd, and 4th moments), power, sample entropy, Hjorth parameters (Activity, Mobility, and Complexity)** are evaluated from various combinations of PRCs.

Mean value of PRCs is evaluated based on the time-domain as follows:

$$Mean\ value = \mu = \frac{1}{N}\sum_{k=0}^{N-1} X[n] \qquad (2.38)$$

Here, $X[n]$ indicates the PRC which is used for analysis and N is the size of the PRC.

Total power and four different higher order moments are evaluated according to the spectrum of signals as follows [96-98]:

$$w_k = \frac{2\pi}{N}k, \quad k = 0,\ 1,\ 2,\ \ldots,\ N-1 \qquad (2.39)$$

$$S(w_k) = \frac{1}{N}|X(w_k)|^2 \qquad (2.40)$$

$$S_T = \sum_{k=0}^{N-1} S(w_k) \qquad (2.41)$$

$$Moment_j = \sum_{k=0}^{N-1}(w_k)^j\, s(w_k), \quad j = 1,\ 2,\ 3,\ 4. \qquad (2.42)$$

where, the PSD of the signal is indicated by $S(w_k)$, and the Discrete Fourier Transform of the signal $X[n]$ is indicated as $X(w_k)$. The size of the related signal is N .

Hjorth parameters are statistical time-domain features and consist of Activity (A_x), Mobility (M_x), and Complexity (C_x) parameters [99]. Activity parameters can be evaluated utilizing the variance of signal amplitude [99, 100].

$$A_x = \big(y(n)\big) = \sigma_y^2 \qquad (2.43)$$

Here, $y[n] = [y_1,\ y_2,\ \ldots,\ y_N]$ and N is the size of the corresponding signal. σ_y indicates the standard deviation of the signal $y[n]$ and it can be evaluated as follows:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}[y[n]-\mu]^2} \tag{2.44}$$

In the above equation, μ indicates the mean value of $y[n]$. Mobility parameter is the ratio of the standard deviations of first-order derivatives. Its mathematical formulation is given as following equation:

$$M_x = \sqrt{\frac{\sigma_{y'}^2}{\sigma_y^2}} = \frac{\sigma_{y'}}{\sigma_y} \tag{2.45}$$

Here, the first-order standard deviation of signal is denoted as $\sigma_{y'}$. Complexity parameter is defined as the ratio between the M_x of the first derivative of the EEG signal and M_x of the EEG signal itself [99, 100].

$$C_x = \frac{M_x\left(y'(t)\right)}{M_x\left(y(t)\right)} = \sqrt{\frac{\frac{\sigma_{y''}^2}{\sigma_{y'}^2}}{\frac{\sigma_{y'}^2}{\sigma_y^2}}} \tag{2.46}$$

Sample entropy is a measurement of time series complexity. The new vector series are defined from the original time series. The sizes of the new vector sequences are defined as m and the size of the original signal is defined as N . The defined m length denotes the embedding dimension [95, 101]. The distance $d\left(u(i),\, u(j)\right)$ between vectors $u(i)$ and $u(j)$ is defined as follows [101]:

$$d\left(u(i),\, u(j)\right) = \max\{|u(i+k)-u(j+k)|\},\ \ 0 \le k \le m-1 \tag{2.47}$$

where, k denotes an index. The probability of ensuring another vector within a distance r from vector is described as [101]:

$$C_i'^m(r) = \frac{1}{N-m+1} \tag{2.48}$$

$$\left\{The\ number\ of\ j,\ j \neq i,\ j \le N-m+1\ such\ that\ d\left(u(i),\, u(j)\right) \le r\right\}$$

$$\emptyset^m(r) = ((N-M+1)^{-1})\sum_{i=1}^{N-m+1} C_i'^m(r) \tag{2.49}$$

Hence, the sample entropy is indicated as [95]:

$$SampEn(m,\ r,\ N) = -\ln\left[\frac{\emptyset'^{m}(r)}{\emptyset'^{m+1}(r)}\right] \tag{2.50}$$

2.2.4 Non-Linear Feature Set

In order to investigate effectiveness of non-linear features, Poincare plot-based features are evaluated for the analysis of different EEG signals. Biomedical signals have non-linear characteristics, so these features are investigated in the literature for different biomedical signals. The Poincare plot measures become an important feature extraction process thanks to its uncomplicated visual explanation and demonstrated clinical ability [53, 54, 102, 103]. These measures can be possible to provide accurate and relevant patterns of physiological signals to provide high classification results for MI task classification.

Poincare plot measures capture the non-linear dynamics embedded in the signal. Poincare plot is a graph or 2D visual representation of each EEG data (x_i) on x-axis and the next EEG data (x_{i+lag}) on the y-axis [54]. These plots are generated from MI EEG signals using determined (x_i), and (x_{i+lag}) intervals of EEG data for each EEG segment. Then, an ellipse is fitted to the generated plot, and the standard deviation of the distance of the points on these graphs denotes the width (SD_1) and length (SD_2) of the fitted ellipse [104]. The detailed mathematical evaluation of Poincare plot measures is given as follows [54, 102, 103]:

$$x_i = \left(x_0,\ x_1,\ \dots,\ x_{N-lag}\right) \tag{2.51}$$

$$x_{i+lag} = \left(x_{lag},\ x_{lag+1},\ \dots,\ x_N\right) \tag{2.52}$$

$$x_w = \frac{x_{i+lag} - x_i}{\sqrt{2}} \tag{2.53}$$

$$x_l = \frac{x_{i+lag} + x_i}{\sqrt{2}} \tag{2.54}$$

$$SD_1 = SD(x_w) \tag{2.55}$$

$$SD_2 = SD(x_l) \tag{2.56}$$

Here, EEG data and the next EEG data intervals are represented with (x_i), and (x_{i+lag}) in Equations (2.51), and (2.52); respectively. According to these intervals (SD_1) and length (SD_2) measurements are calculated as features. SD denotes the standard deviation of the defined time interval vectors in in Equations (2.55), and (2.56). SD_1, and SD_2 measures are evaluated based on the defined lag value. In the literature, the commonly utilized lag value is 1 [53, 54, 105, 106]. In this thesis, we aimed to investigate the effectiveness of various lag values which are from 1 to 10 for extremity movement task classification studies. In addition to the evaluation of SD_1 and SD_2, the product ($SD_1 SD_2$) and the ratio ($\frac{SD_1}{SD_2}$) are calculated to investigate the relation between these measures. Therefore, a total of four non-linear features are extracted for each EEG signal which in the defined a lag value. The effectiveness of different lag values is investigated generating ten different feature sets, separately. In this thesis, Poincare features are evaluated where only $lag = 1$ for the investigation of the non-linear feature set in finger movement classification.

2.3 Feature Selection

Feature selection process can be optionally applied to decrease complexity of classification process and improve classifier performance for selection or reduction of effective features from all features in feature sets after feature extraction step [54, 107, 108]. In this thesis, statistical significance-based feature selection methods such as ANOVA and independent t-test are applied to improve classifier performance selecting relevant and discriminative EEG features. In addition to statistical significance-based feature selection method, PCA-based feature selection method is performed in order to examine comparatively the effectiveness of ANOVA in only finger movement classification analysis.

2.3.1 Statistical Significance

In this thesis, the statistical significance-based feature selection method is performed to select effective feature combinations in different feature sets in our EEG signal processing. These tests are performed for each provided feature sets separately. According to the class number of MI tasks, the type of statistically significant-based feature selection methods is defined. In this thesis, binary and multi-class classification models are performed using various feature sets [109]. Therefore, two different types of statistically significant-based feature selection methods such as ANOVA test and the independent t-test applied to select features in multi-class and binary-class classifications, respectively. In binary classification models, the independent t-test which is widely used to indicate the significance of differences between features of two different classes is performed for reduction of features [53]. In multi-class classification models, ANOVA test which is used to check that there is a significant difference between features of multiple classes is performed to select features [40, 54, 109]. According to these tests, the statistical significance of all extracted features in feature sets is defined evaluating p values. The indicated statistical significance level (α) is 0.05. The features that obtain this range are determined as statistically significant and selected features.

2.3.2 Principal Component Analysis

PCA-based feature selection, which is an effective feature reduction method, is performed in finger movement classification in order to compare effectiveness of ANOVA-based feature selection. It is known as a multivariate statistical transformation technique to remove similarity between features. The linearly-independent perpendicular features are generated using PCA. The number of them indicates the system parameter, covering the percentage ratio of the variance of the initial variables. Each of new variables is defined as the principal component [107, 110, 111].

PCA provides the principal components of the data based on the evaluation of eigenvalue and eigenvector of the covariance matrix after data normalization. X is a matrix with size of nxm d i th row of it with size of m defined as $(X_i,\ i = 1,\ 2,\ \dots,\ n)$. The covariance matrix is evaluated using mean value of the data as follows:

$$C = \sum_{i=1}^{n}(x-\mu)(x-\mu)^T \tag{2.57}$$

where, μ is the mean value of the data and C indicates the covariance matrix. The eigenvalues (λ) and eigenvectors (V) of C is evaluated as follows:

$$\det(\lambda I - C) = 0,\ \ (\lambda_k I - C)xV_k = 0 \tag{2.58}$$

The eigenvalues are listed in ascending series and the eigenvectors matched with the largest eigenvalues are found. The selected data is generated with projection of normalized data onto K eigenvectors [107, 110].

2.4 Classification

In our thesis, the extracted different feature sets are calculated using eight different classifiers such as DT, LDA, NB, k-NN, EL, NNs and KA to distinguish different MI tasks. In addition to these classifiers, LR is also used to classify MI EEE segments in binary-class classifications. The fundamental information about these classifiers is available in below:

2.4.1 Decision Tree

DT is a machine learning algorithm which can divide the data into different classes. It can be applied for both classification and regression analysis. The branch and nodes which are in this algorithm are likened to tree-like structures and they give the name of algorithm. Training in this algorithm is performed based on the order of decision rules. If a decision is completed a leaf node is generated while when a decision is not completed a decision node that is different branch is generated. In this thesis, three different tree algorithms such as fine, medium, and coarse are performed for the classification process [112].

2.4.2 Discriminant Analysis

DA classification aims to separate the independent variables in the data accurately into homogeneous groups [113]. LDA among these algorithms indicates group elements and evaluates the probability of characterizing different groups for each element. The group which obtains the highest probability score is indicated as the predicted group of elements. It generates a linear discrimination function. In this algorithm, the predictors are accepted to be normally distributed (Gauss distribution). And also, it assumes that different classes have class specified elements and equal variance/covariance. However, the variance/covariance equality is not accepted in QDA algorithm. It assumes that covariance matrix can be divergent for each class. Hence, it generates a second order discrimination function in process [114, 115].

2.4.3 Naïve Bayes

NB is a probabilistic machine learning algorithm using variables' independence and normalcy and Bayes theorem that classification is applied in accordance with probability basics. The calculation of the membership probability of a sample to all class in feature set is fundamental process of this algorithm.

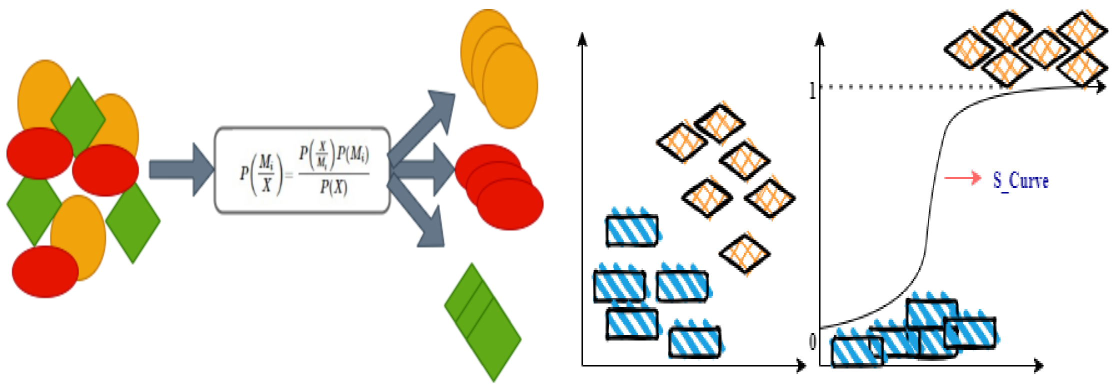
A sample X in the feature set is defined as $X = \{x_1, x_2, \ldots, x_n\}$ and n is the number of features. Classes in feature set are defined as $\{M_1, M_2, \ldots, M_m\}$ and m indicates the number of classes. The probability that each X data in data set is a member of the M_i class is evaluated as:

$$P\left(\frac{M_i}{X}\right) = \frac{P\left(\frac{X}{M_i}\right)P(M_i)}{P(X)},\ if\ P\left(\frac{M_i}{X}\right) > P\left(\frac{M_j}{X}\right),\ 1 \leq j \leq m,\ j \neq i \qquad (2.59)$$

Therefore, the highest probability of membership defines the class of the data. According to this formula, X data is labeled to the M_i. The class prior probabilities are represented by $P(M_i)$, the prior probability of X sample is represented with $P(X)$. The probability of X conditioned on M_i is denoted as $P\left(\frac{X}{M_i}\right)$ and $P\left(\frac{M_i}{X}\right)$ denotes the probability of M_i conditioned on X [112, 116]. Medium and Gaussian NB algorithms are used for classification process in this thesis. The basic scheme of NB classification process is given in Figure 2.2a.

2.4.4 Logistic Regression

LR is a commonly applied statistical machine learning process in which binary classification results are generated such as yes/no, 1/0. It is related to a set of independent variables as given in following equation:



(a) NB classifier simulation (b) LR classifier simulation

Figure 2.2: The basic scheme of the (a) NB classifier, (b) LR classifier.

$$Logit(P_1) = \ln\left(\frac{P_1}{1-P_1}\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n \tag{2.60}$$

Here, P_1 is probability of an event, β_0 indicates the intercept, $\{\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n\}$ indicates the coefficients related to the independent variables $\{X_1 + X_2 + \cdots + X_n\}$. In process of LR algorithm, maximum likelihood estimation is mainly utilized to evaluate the coefficients. The probability of an event as a logistic function of the independent variables is non-linear as shown in following equation:

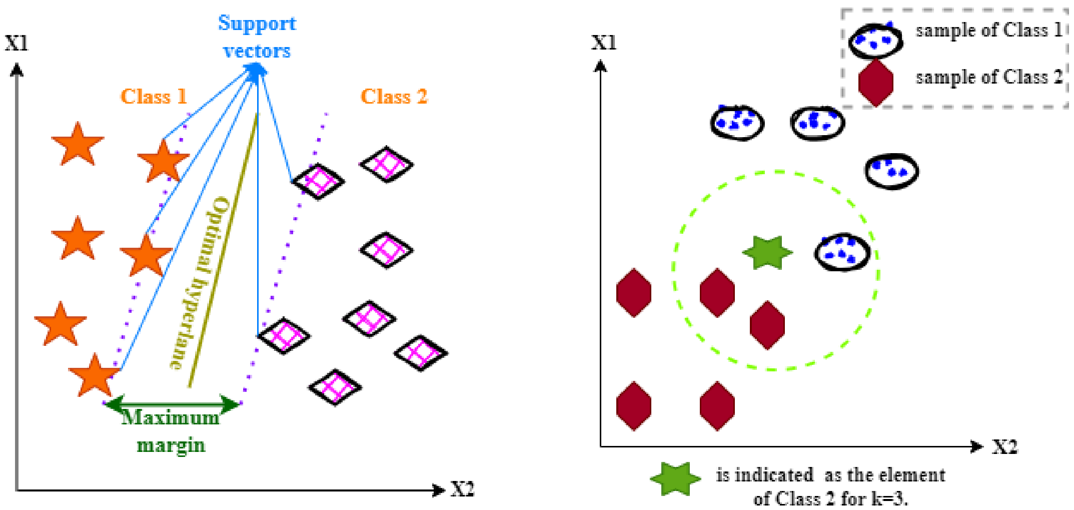
$$P_1(x) = \frac{P_1}{1+e^{-Logit(P_1(x))}} \tag{2.61}$$

where, P_1 is defined as probability value and takes between 0 and 1. When the result of $P_1(x)$ equation is $-\infty$, $P_1 = 0$, and when $P_1(x) = \infty$, the probability equals 1. The basic scheme of classification process of LR is represented in Figure 2.2b.

2.4.5 Support Vector Machine

SVM is a successful machine learning algorithm and utilized in both classification and regression analysis. It classifies the data based on the geometric characteristic of this data. Firstly, the elements of the dataset which consists of n features are settled as the elements of the coordinate system that is n-dimensional space. Then, the classification is carried out based on obtaining the hyperplane that discriminates the classes best. Different hyperplanes can be constructed for discrimination of two classes. However, the selection of the hyperplane that the highest and accurate classification performance may be provided is crucial from the different hyperplanes. Let, (x_n, y_n) is defined as a linearly separable sample example. n denotes the size of the feature set and y which takes value of –1 or 1 denotes as class label. The hyperplane can be formulated as $f(x) = wx + b$ here w and b denote the hyperplane parameters and the offset, respectively. The main aim here is to provide the maximum margin. The dashed lines (represented in Figure 2.3a) indicate the decision boundaries which are placed on support vectors. The margin is defined as the distance between these support vectors which belongs to two different classes. Therefore, the data placed on different sides of the optimal hyperplane is defined as a sample of the different class [107, 113, 117].

The basic binary-class SVM classification is shown in Figure 2.3a. Six different algorithms of SVM classifier such are Linear, Quadratic, Cubic, Fine Gaussian, Medium Gaussian, and Coarse Gausssian are used for classifications in this thesis.



(a) SVM classifier simulation (b)kNN classifier simulation

Figure 2.3: The basic scheme of the (a) SVM classifier, (b) kNN classifier.

2.4.6 K-Nearest Neighbour

k-NN, is a learning-based machine learning algorithm, evaluates the closeness of new data with defined classes. The distance of new sample and all the data in the feature set is evaluated. The closeness of new data is examined checking k nearest neighbor and minimum distance is evaluated. Finally, whichever class has the most elements among the determined neighbors is labeled as the class of the new sample. Different distance measurements methods such as Euclidean, Manhattan, Minkowski, and Hamming can be applied to calculate distance. Among these methods, the mostly applied method is Euclidean distance and its formulation is given in following equation [54, 107]:

$$Euclidean = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{2.62}$$

In this thesis, Fine, Medium, Coarse, Cubic, Cosine, and Weighted algorithms of the k-NN classifier were performed for classifications. The process of k-NN algorithm for binary classification is basically represented in Figure 2.3b.

2.4.7 Ensemble Learning

EL is defined as meta-algorithms which merges multiple pattern recognition techniques into a single discrimination model (classifier) to enhance deviation (boosting), an/or predictions (stacking) and lessen variance (bagging) [107, 114, 118-120]. This classifier assumes that single classifiers mostly cannot provide a specific and accurate classification accuracy owing to possible noise in the data, overlapping data distributions, and outliers. The EL algorithm is generated and used for classifications considering there is no single classifier that performs best for all classification tasks. Hence, it is necessary to use EL algorithms in some classifications. An EL algorithms can be mostly created in two different methods such as sequential ensemble learning methods (AdaBoost) and parallel ensemble learning methods (RF). In this thesis, Boosted, Bagged, Subspace Discriminant, Subspace k-NN, and RUSBoosted Trees which are introduced in the process of EL classifiers are applied.

2.4.8 Kernel Approximation

KA algorithms can be utilized to perform non-linear classification of data including many samples [121, 122]. In large datasets, KA classifiers are inclined to train and predict faster than SVM algorithms accompanied by Gaussian kernels [122]. Gaussian kernel algorithms plan predictors in a low-dimensional space to high-dimensional space. Then, the linear model is constructed to convert predictors in a high-dimensional area [121, 122]. In this thesis, SVM and LR KA algorithms are used to classify.

2.4.9 Neural Networks

NN algorithms, are complex classification algorithms, mimic the human brain and provide accurate classification accuracy. Its deep neural structure which includes the number of layers and different parameters make the training process longer [123, 124]. NNs structures include three fundamental structures such as input layer, fully connected layers, and output layer as shown in Figure 2.4.

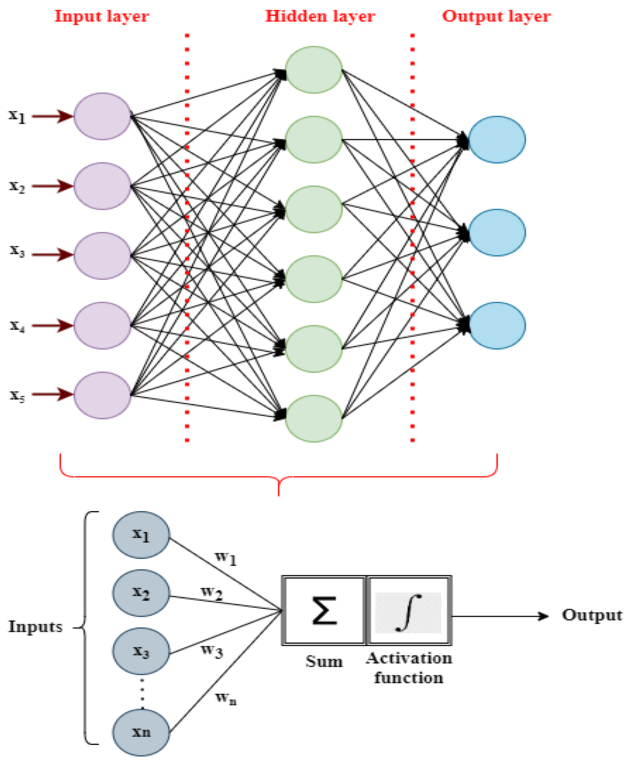


Figure 2.4: The basic representation of NNs structures.

Various NN algorithms are available and the number of fully connected layers between the input and output layers may differ in different algorithms. The number of fully connected layers determines the complexity of the classifier. When the size and number of these layers increase, the complexity of the model is also increased [123, 124]. The first fully connected layer of the NN has a relationship from the network input (predictor data), and each subsequent layer has a relationship from the preceding layer. In fully connected layers the input data (features) and a weight matrix are multiplied between each other and a bias vector is included into this evaluation as shown in Figure 2.4. An activation function accompanies each fully connected layer. Finally, the final fully connected layer and the following softmax function give the output of NN as classification scores and prediction labels [125-127]. In this thesis, Narrow, Medium, Wide, Bi-layered, and Tri-layered NN algorithms are used for classifications. The basic representation of the NN algorithm is given in Figure 2.4.

2.5 Performance Evaluation

In this thesis, ACC is defined as performance metric and is used to evaluate performances of different classifiers. Accuracy is the ratio of the total number of true predictions and is evaluated based on the confusion matrix. The confusion matrix represents the number of true and false predictions performed by classifier compared to real labels in the data. The confusion matrix with nxn is created according to number of classes (n). Accuracy metric formulation is given in below:

$$Accuracy\ (ACC) = \frac{TP+TN}{TP+FP+TN+FN} x100\% \tag{2.63}$$

Here, TP, is actual true, predicted as true by model TN, actual false, predicted as false by model. While FP is not actually in class true but is predicted in class false by model, FN is actually in class true but is predicted in class false.

K-fold CV method is utilized to show performances of the classification algorithms [87, 107, 112]. K is defined as 5 in our experimental analysis. In our analysis, firstly feature set is divided into train (80% samples of feature sets) and test set (20% samples of feature sets). Then, 5-fold CV method is applied in training feature set to provide a consistent accuracy for training process. In this process, training feature set is separated into 5 equal size subsets. The classification process repeated 5 times, and each time 4/5$\left(\frac{K-1}{K}\right)$ of the subsets are used for training, and the remaining 1/5$\left(\frac{1}{K}\right)$is utilized for validation. Finally, the average performance of 5 (K) is evaluated as training accuracy value. The proposed model is also tested using test data and its performance is evaluated based on the accuracy metric. The basic representation of 5-fold cross-validation is given in Figure 2.5.

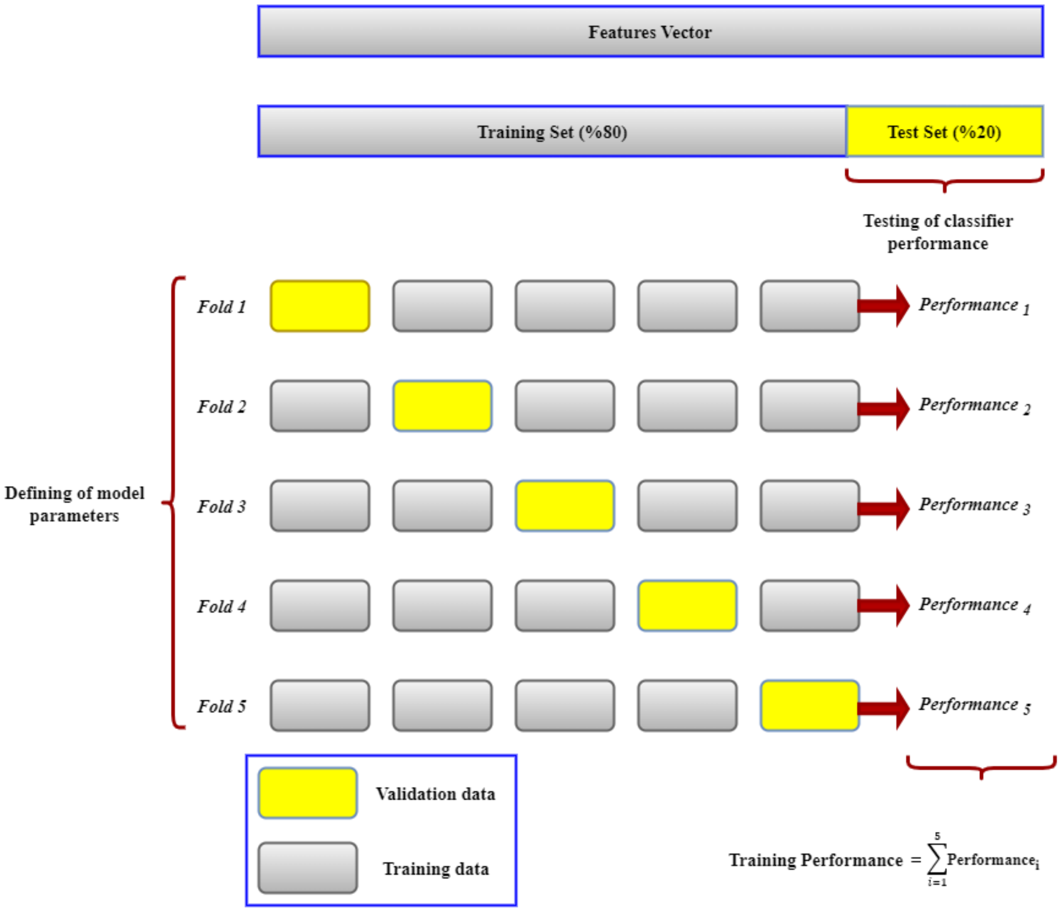


Figure 2.5: The representation of 5-fold cross-validation process-based classification used in our studies.

Chapter 3

Classification of Extremity Movement Task

In this thesis, four different feature sets are presented to distinguish MI task EEG segments. These are time-domain, frequency-domain, time-frequency domain and non-linear feature sets. Additionally, two different combination feature sets are created using different feature extraction approaches and the effectiveness of combination of different features is investigated. In addition, the statistically significance-based feature selection methods are applied and effectiveness of them is investigated in both binary-class and multi-class extremity movement classification. Finally, the results of these six different feature sets and effectiveness of statistically significance-based feature selection methods are compared in line with the classification performances of different machine learning algorithms utilized in our thesis.

3.1 Experimental Data set

In this thesis, the binary-class and multi-class extremity movement task classification analyzes are performed using BCI Competition IV Dataset IIa. The 22 EEG signals which belong to four different MI tasks are supplied from 9 subjects in this dataset. MI tasks are the imagination of the right hand, left hand, both feet, and tongue. In binary classification studies, we used right and left hands MI task EEG signals while in multi-class extremity movement classification studies, all of MI task EEG signals were used. Firstly, the MI task EEG segments performed during 3 sec for each trial are decomposed from EEG time series in preprocessing of EEG signals. Then, all channels of EEG signals were used to extract features for each 3 sec EEG segment. The binary-class extremity movement classification studies performed in this thesis are represented in Figure 3.1.

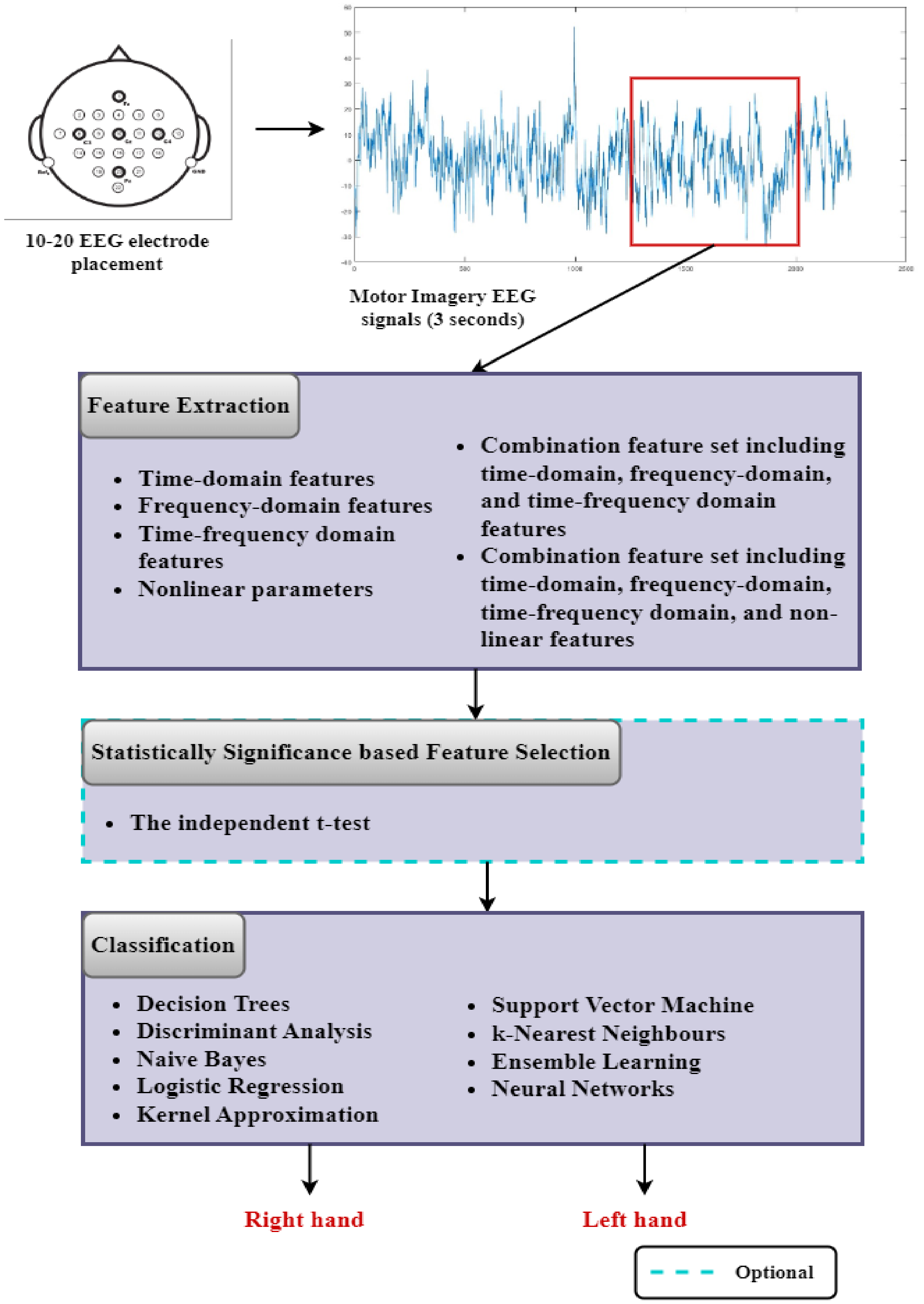


Figure 3.1: The block diagram of the proposed binary-class extremity movement classification approach.

The multi-class extremity movement classification studies performed in this thesis are represented in Figure 3.2.

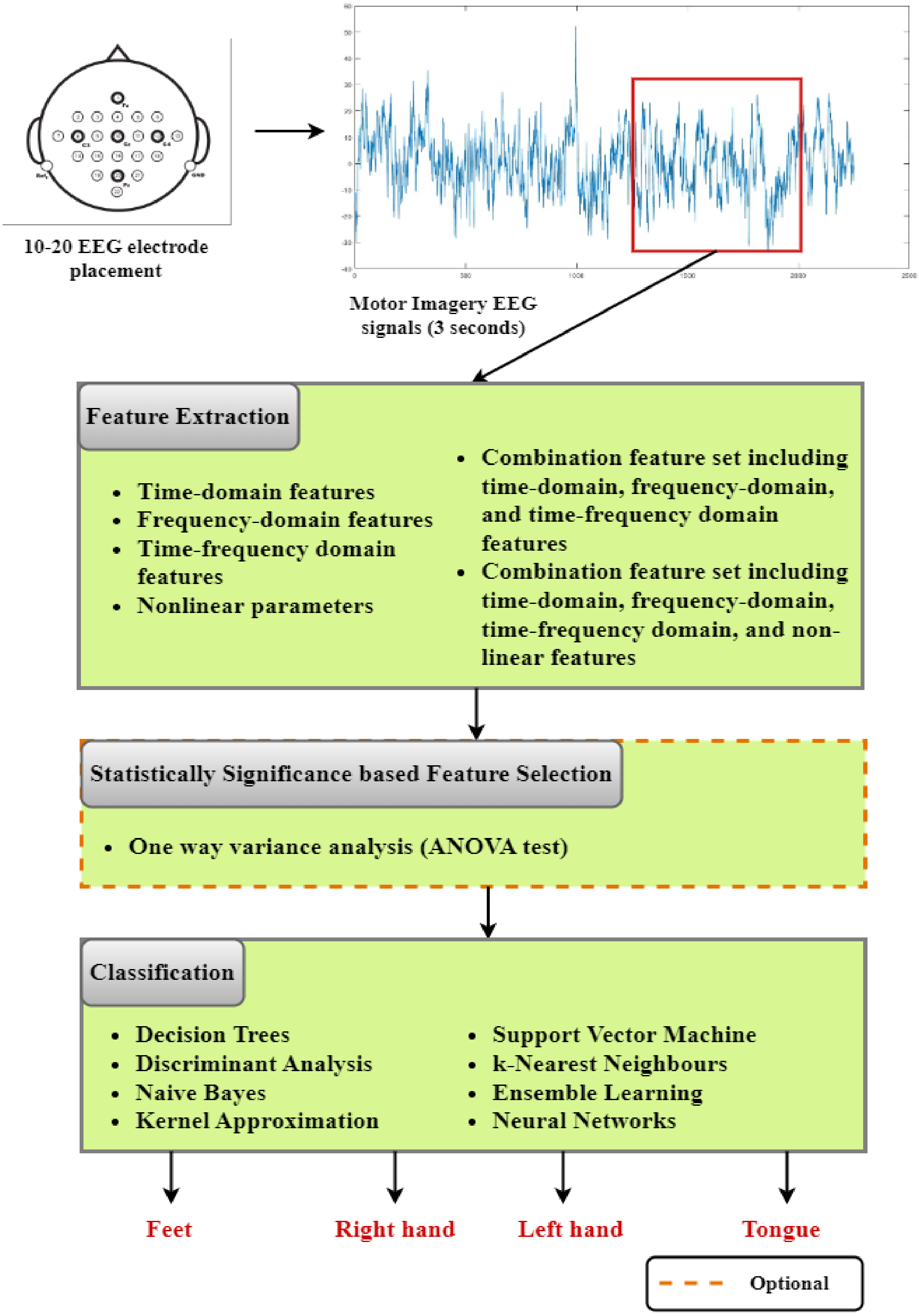


Figure 3.2: The block diagram of the proposed multi-class extremity movement classification approach.

3.2 Feature Extraction

24 time-domain, 15 frequency-domain, 15 time-frequency domain, and 4 non-linear features are evaluated for each EEG channel of each EEG segment. In addition to them, the combinations sets are generated to improve classifier performances. These feature sets and their different combination feature sets are analyzed using various machine learning algorithms for binary-class and multi-class extremity movement classifications, separately.

3.2.1 Time-domain Feature Set

After the extraction of EEG segments, the time-domain feature set was obtained evaluating 24 different amplitude and statistical information-based features in the time-domain. The mathematical formulations of these features are represented in Section 2.2.1. In the time-domain feature extraction-based approach, a total of 2592x528 and 1296x528 size time-domain feature sets are created for multi-class and binary-class extremity movement task classifications, respectively.

3.2.2 Frequency-domain Feature Set

To create the frequency-domain feature set, the frequency domain of EEG segments is obtained using FT and five different EEG sub-bands are decomposed for each EEG segment. Energy, variance, and entropy measures are evaluated using EEG sub-bands. The mathematical formulations of energy, variance and entropy values are represented in Section 2.2.2. In the frequency-domain feature extraction-based approach, a total of 2592x330 and 1296x330 size frequency-domain feature sets are created for multi-class and binary-class extremity movement task classifications, respectively.

3.2.3 Time-Frequency Domain Feature Set

To create the frequency-domain feature set, the time-frequency domain of EEG segments is obtained using WT and five different EEG sub-bands are decomposed for each EEG segment. Energy, variance, and entropy measures are evaluated using EEG sub-bands. Haar mother wavelet and 7 level sub-band decomposition are utilized for our analysis. The mathematical formulations of energy, entropy, and variance values based on the WT methos are represented in Section 2.2.3.

In the time-frequency domain feature extraction-based approach, a total of 2592x330 and 1296x330 size time-frequency domain feature sets are created for multi-class and binary-class extremity movement task classifications, respectively.

3.2.4 Non-linear Feature Set

Non-linear feature sets are created using Poincare plot-based measures. 4 different non-linear features are evaluated to supply information about the non-linear dynamics embedded in EEG signals for each EEG segment where $lag = m$. In this thesis, m defined from 1 to 10 and a non-linear feature set is created for each m value to investigate the effectiveness of different m values, separately. Poincare plot measures' formulations are given in Section 2.2.4.

In the non-linear feature extraction-based approach, a total of 2592x88 size 10 different non-linear feature sets are created for multi-class extremity movement task classification. Additionally, a total of 2592x880 size combination of 10 non-linear feature sets is created for analysis.

3.2.5 Combination Feature Set Including Time-domain, Frequency-domain, and Time-frequency domain Features

In addition to four different feature sets, the effectiveness of the combination of different feature sets are investigated to improve the classification performance in our thesis studies. Combination feature sets are created using 24 time-domain, 15 frequency-domain, and 15 time-frequency domain features for each EEG channel of each EEG segment. In the combination feature set-based analysis, a total of 2592x1188 and 1296x1188 size the combination feature sets are created for multi-class and binary-class extremity movement task classifications, respectively.

3.2.6 Combination Feature Set Including Time-domain, Frequency-domain, Time-frequency domain, and Non-linear Features

In our second combination feature sets, we added 8 non-linear features into our first combination feature set which is created using 24 time-domain, 15 frequency-domain, and 15 time-frequency domain features for each EEG channel of each EEG segment. 4 non-linear features are evaluated for each EEG segments where 2 different lag conditions which are $lag = 1$ and $lag = 9$. A total of 176 non-linear features are evaluated from all EEG channels of each EEG segment based on the 2 different lag conditions and added to our previous combination sets. In our second combination feature set-based analysis, a total of 2592x1364 and 1296x1364 size the combination feature sets are created for multi-class and binary-class extremity movement task classifications, respectively.

3.3 Feature Selection Using Statistically Significance

In our thesis, the statistically significance-based feature selection method is presented. The effectiveness of this method is investigated in six different feature sets for both multi-class and binary-class extremity movement task classifications.

Table 3.1: Sizes of all feature sets and t-test selected feature sets used in binary-class extremity movement task classification.

Feature Set	**All Features**	**T-test Selected Features**
TD	(1296x528)	(1296x44)
FD	(1296x330)	(1296x28)
WT	(1296x330)	(1296x13)
TD+FD+WT	(1296x1188)	(1296x85)
TD+FT+WT+P	(1296x1364)	(1296x91)

The statistically significance-based feature selection methods which are the independent t-test and ANOVA test are used to select statistically significant features in feature sets for binary-class and multi-class extremity movement task classifications, respectively. Table 3.1 represents the number of all features and t-test selected features in six different feature sets which are used in our binary-class extremity movement task classifications. Table 3.2 represents the number of all features and ANOVA selected features in six different feature sets which are used in our multi-class extremity movement task classifications.

Table 3.2: Sizes of all feature sets and ANOVA selected feature sets used in multi-class extremity movement task classification.

Feature Set	All Features	ANOVA Selected Features
TD	(2592x528)	(2592x345)
FD	(2592x330)	(2592x102)
WT	(2592x330)	(2592x104)
TD+FD+WT	(2592x1188)	(2592x551)
TD+FT+WT+P	(2592x1364)	(2592x612)

3.4 Results and Discussions of Binary-Class Extremity Movement Task Classification

EEG signals including MI tasks provided from 22-channel EEG recordings of 9 subjects were analyzed utilizing different feature sets and various classifiers. The different features extraction approaches including time-domain, frequency-domain, time-frequency domain, and non-linear features were performed for binary-class extremity movement task (right hand and left hand MI tasks) classification after obtaining of 3 sec MI EEG segments. Time-domain (24 different statistical and amplitude-based measures), frequency-domain (energy, variance, and entropy measures of FT-based five different EEG sub-bands), time-frequency domain (energy, variance, and entropy measures of WT-based five different EEG sub-bands), and non-linear (4 different Poincare plot measures) feature sets were created from 22-channel EEG signals. The effectiveness of these four different feature sets and their two different combination feature sets are investigated, separately. In addition, the effectiveness of the independent t-test based feature selection process is investigated with all extracted feature sets. Finally, DT, DA, NB, LR, SVM, k-NN, EL, NNs, and KA machine learning algorithms are performed for classification, and the results are evaluated. All signal processing (signal segmentation, feature extraction, feature selection, and classification) and performance analyzes were implemented using MATLAB software. The performances of these six different features sets are compared using 9 different classifiers.

Performance evaluation results of our proposed approach including different feature sets, the independent t-test based feature selection, and various classifiers are given in Tables 3.3-3.8. In these tables TD, FD, WT, and P indicate that the features for classifications using the time-domain, frequency-domain, time-frequency domain, and non-linear information, respectively. On the other hand, TD+T-test, FD +T-test, and WT+T-test indicate that the independent t-test selected statistically significant features for classifications using the time-domain, frequency-domain, and time-frequency domain, respectively. The classifications performed using the first combination feature set including time-domain, frequency-domain, and time-frequency domain features are indicated as TD+FD+WT and the classifications performed using the independent t-test selected statistically significant features of this combination set are indicated as TD+FD+WT+T-test. The classifications performed using the second combination feature set including time-domain, frequency-domain, time-frequency domain, and non-linear features are indicated as TD+FD+WT+P and the classifications performed using the independent t-test selected statistically significant features of this combination set are indicated as TD+FD+WT+P+T-test. The boldface characters in table cells represent the best classification performance for each approach and classification algorithms (in Tables 3.3-3.8).

The performance evaluation of all time-domain features and the t-test selected statistically significant time-domain features for binary-classification is summarized in Table 3.3. We obtained the highest accuracy value of 61.26% using all time-domain features evaluated from EEG segments and EL algorithm while the NB algorithm performed the worst accuracy value of 52.62% for the same features. When the t-test selected statistically significant time-domain features evaluated using various classifiers, we achieved the highest 56.64% classification accuracy using LR algorithm and the worst accuracy value of 51.08% using KA algorithm. To discover the effectiveness of the independent t-test selection process, we analyzed and compared performance results of TD and TD+T-test classifications. It was observed that t-test based feature selection diminished classifier performances in all classifiers except two (NB and k-NN). Results of all classification using time-domain and statistically significant time-domain features are given in Table 3.3.

Table 3.3: Performance results (%) for binary-class extremity movement task classification using time-domain feature set.

Models	**Accuracy**	
	TD	**TD+T-test**
Decision Tree	56.56	55.02
Discriminant Analysis	57.64	56.02
Logistic Regression	56.17	55.79
Naive Bayes	52.62	55.17
Support Vector Machine	59.57	**56.64**
k-Nearest Neighbours	53.24	54.32
Ensemble Learning	**61.26**	57.72
Neural Networks	58.72	53.01
Kernel Approximation	54.24	51.08

The performance evaluation of all frequency-domain features and the t-test selected statistically significant frequency-domain features for binary-classification is summarized in Table 3.4. We obtained the highest accuracy value of 60.03% using all frequency-domain features evaluated from EEG segments and EL algorithm while the LR and k-NN algorithms performed the worst accuracy value of 52.01% for the same features. When the t-test selected statistically significant frequency-domain features evaluated using various classifiers, we achieved the highest 61.34% classification accuracy using LR algorithm and the worst accuracy value of 53.55% using KA algorithm. To discover the effectiveness of the independent t-test selection process in frequency-domain feature set, we analyzed and compared performance results of FD and FD+T-test classifications. It was observed that t-test based feature selection improved classifier performances in all classifiers except two (DT and KA). Results of all classification using frequency-domain and statistically significant frequency-domain features are given in Table 3.4.

Table 3.4: Performance results (%) for binary-class extremity movement task classification using frequency-domain feature set.

Models	Accuracy	
	FD	**FD+T-test**
Decision Tree	57.56	57.48
Discriminant Analysis	53.86	61.11
Logistic Regression	54.63	**61.34**
Naive Bayes	52.01	55.79
Support Vector Machine	55.63	59.03
k-Nearest Neighbours	52.01	54.78
Ensemble Learning	**60.03**	60.26
Neural Networks	56.48	57.18
Kernel Approximation	55.94	53.55

The performance evaluation of all time-frequency domain features and the t-test selected statistically significant time-frequency-domain features for binary-classification is summarized in Table 3.5. We obtained the highest accuracy value of 52.70% using all time-frequency domain features evaluated from EEG segments and DT algorithm while the LR and k-NN algorithms performed the worst accuracy value of 49.85% for the same features. In DA, SVM, NN, and KA, we did not perform classification using all time-frequency domain features, because the proposed feature set is not suitable for the classifier structure. When the t-test selected statistically significant time-frequency domain features evaluated using various classifiers, we achieved the highest 54.71% classification accuracy using LR algorithm and the worst accuracy value of 49.15% using KA algorithm. To discover the effectiveness of the independent t-test selection process in time-frequency domain feature set, we analyzed and compared performance results of WT and WT+T-test classifications. When the performed classifications were examined for the case where all features were used, it was observed that the independent t-test based feature selection method increased the performance in all classifiers except one out of 5 classifiers. Results of all classification using time-frequency domain and statistically significant time-frequency domain features are given in Table 3.5.

Table 3.5: Performance results (%) for binary-class extremity movement task classification using the time-frequency domain feature set.

	Accuracy	
Models	**WT**	**WT+T-test**
Decision Tree	**52.70**	50.62
Discriminant Analysis	N/A	50.93
Logistic Regression	49.85	50.77
Naive Bayes	51.16	**54.71**
Support Vector Machine	N/A	50.93
k-Nearest Neighbours	49.85	50.69
Ensemble Learning	51.39	53.94
Neural Networks	N/A	50.54
Kernel Approximation	N/A	49.15

The performance evaluation of all non-linear feature sets with various classifiers is summarized in Table 3.6. In this table, Lag (1)-Lag (10) indicate that the features for classifications are evaluated by using the corresponding lag value. Additionally, “All lags” indicates that the classifications are carried out using the combination feature set provided by combining the features from 10 different lag values. The results revealed that the non-linear feature set extracted for lag=6 condition achieved the highest accuracy value of 63.35% using DA classifier and the worst accuracy value of 48.53% is evaluated using All lags feature set and NB classifier. On the other hand, we analyzed the effectiveness of different lag values for 9 classifiers. We observed that the higher classification performance is obtained in 2 (NN and KA) classifiers using feature set for lag=4 condition, in 3 (DA, LR, and EL) classifiers using feature set for lag=6 condition, and in 1 classifier using feature sets for lag=7, lag=9, lag=10, and All lags conditions. Among all non-linear feature sets, the most successful feature set is evaluated as the 6th feature set generated where lag=6.

Table 3.6: Performance results (%) for binary-class extremity movement task classification using the non-linear feature set.

Model	Lag (1)	Lag (2)	Lag (3)	Lag (4)	Lag (5)	Lag (6)	Lag (7)	Lag (8)	Lag (9)	Lag (10)	All lags
Decision Tree	52.62	52.01	52.70	54.63	52.70	54.71	**55.40**	54.24	54.32	53.94	53.32
Discriminant Analysis	59.03	60.34	61.73	58.80	60.11	**63.35**	60.57	59.49	59.95	60.03	54.63
Logistic Regression	59.03	60.49	62.35	59.72	61.03	**62.65**	61.11	60.57	60.57	60.42	53.24
Naive Bayes	49.85	49.07	49.77	49.61	49.07	50.46	50.08	50.31	51.93	**52.24**	48.53
Support Vector Machine	58.64	61.57	62.58	62.35	61.73	62.19	62.35	61.57	60.80	60.03	**63.04**
k-Nearest Neighbours	51.00	52.47	53.01	53.78	53.86	53.86	55.09	53.32	**55.71**	53.86	53.16
Ensemble Learning	58.95	60.26	61.27	61.42	61.34	**62.81**	61.81	60.03	60.42	58.80	61.19
Neural Networks	60.42	58.64	59.65	**61.88**	58.64	60.12	60.88	59.57	59.26	57.02	60.57
Kernel Approximation	52.16	53.70	52.93	**54.63**	52.39	54.17	52.24	53.16	52..85	54.17	53.86

In order to compare the effectiveness of different feature sets, the classification is carried out with the combination of time-domain, frequency-domain, and time-frequency domain features. The performance evaluation results of this combination feature set and the independent t-test selected feature set from the combined feature set are summarized in Table 3.7. In classification performed using combined feature set, the EL algorithm obtained the maximum accuracy (58.10%) and k-NN obtained the worst accuracy (49.85%) using same features. In DA, SVM, NN, and KA, we did not perform classification using all combination set features, because the proposed feature set is not suitable for the classifier structure. On the other hand, in the classifications performed using the independent t-test selected feature set, the EL algorithm provided the maximum accuracy (62.96%) and KA provided the worst accuracy (50.00%) using same features. To discover the effectiveness of the independent t-test selection process in the combination feature set, we analyzed and compared performance results of TD+FD+WT and TD+FD+WT+T-test classifications. When the performed classifications were examined for the case where all features were used, it was observed that the independent t-test based feature selection method increased the performance in all classifiers except one out of 5 classifiers. Results of all classification using the combination set and the selected statistically significant combination set features are given in Table 3.7.

Table 3.7: Performance results (%) for binary-class extremity movement task classification using the combined (TD+FD+WT) feature set.

	Accuracy	
Models	**TD+FD+WT**	**TD+FD+WT+T-test**
Decision Tree	56.71	55.25
Discriminant Analysis	N/A	51.23
Logistic Regression	49.92	51.00
Naive Bayes	53.47	57.02
Support Vector Machine	N/A	51.16
k-Nearest Neighbours	49.85	50.31
Ensemble Learning	**58.10**	**62.96**
Neural Networks	N/A	50.85
Kernel Approximation	N/A	50.00

The classification is performed in our second combination set including time-domain, frequency-domain, time-frequency domain, and non-linear features. We added non-linear feature sets extracted for lag=1 and lag=9 conditions, into our previous combination set. The performance evaluation results of this combination feature set and the independent t-test selected feature set from this combined feature set are summarized in Table 3.8. In classification performed using combined feature set, the EL algorithm obtained the maximum accuracy (57.30%) and NB obtained the worst accuracy (48.50%) using same features. On the other hand, in the classifications performed using the independent t-test selected feature set, the EL algorithm provided the maximum accuracy (61.86%) and KA provided the worst accuracy (49.92%) using same features. To discover the effectiveness of the independent t-test selection process in the combination feature set, we analyzed and compared performance results of TD+FD+WT+P and TD+FD+WT+P+T-test classifications. When the performed classifications were examined for the case where all features were used, it was observed that the independent t-test based feature selection method increased the performance in all classifiers except one out of 5 classifiers. Results of all classification using our second combination set and the selected statistically significant combination set features are given in Table 3.8.

Table 3.8: Performance results (%) for binary-class extremity movement task classification using the combined (TD+FD+WT+P) feature set.

Models	Accuracy	
	TD+FD+WT+P	**TD+FD+WT+P+T-test**
Decision Tree	56.60	56.50
Discriminant Analysis	N/A	52.10
Logistic Regression	49.90	51.10
Naive Bayes	48.50	57.10
Support Vector Machine	N/A	51.40
k-Nearest Neighbours	49.80	50.80
Ensemble Learning	**57.30**	**61.86**
Neural Networks	N/A	50.54
Kernel Approximation	N/A	49.92

Also, the independent t-test selected features in four different sets are analyzed to investigate effects of different features and EEG channels. Firstly, we investigated the independent t-test selected time-domain features. The list of 24 different time-domain features with their abbreviations are available in Table 3.9. Channel-based t-test selected statistically significant time-domain feature distribution is given in Table 3.10. A total of 44 time-domain features were indicated as statistically-significant features using t-test. Among the 24 different time-domain features, the selected features are maximum value (in 3 EEG channels), mean value (in 12 EEG channels), kurtosis (in 4 EEG channels), skewness (in 3 EEG channels), Q1 (in 5 EEG channels), Q2 (in 13 EEG channels), Q3 (in an EEG channel), and slope-change value (in 3 EEG channel). Among 22 EEG channels, more statistically significant features were selected from some channels (6^{th}, 12^{th}, 13^{th}, and 18^{th} EEG channels). Also, the statistically significant features were not selected from some channels as can be seen from Table. As a result, it was observed that feature selections were made from certain channels and certain features with the t test. However, it has been observed that the independent t-test generally cannot improve the classifier performance in the time domain feature set.

Table 3.9: Time-domain features.

Time-domain features			
$\mathbf{T_1}$	Minumum value	$\mathbf{T_{13}}$	Kurtosis
$\mathbf{T_2}$	Maximum value	$\mathbf{T_{14}}$	Skewness
$\mathbf{T_3}$	Mean	$\mathbf{T_{15}}$	Hjorth parameters (Activity)
$\mathbf{T_4}$	Standard deviation value	$\mathbf{T_{16}}$	Hjorth parameters (Mobility)
$\mathbf{T_5}$	Integrated EEG value	$\mathbf{T_{17}}$	Hjorth parameters (Complexity)
$\mathbf{T_6}$	Mean absolute value	$\mathbf{T_{18}}$	Signal range
$\mathbf{T_7}$	Simple square integral	$\mathbf{T_{19}}$	First inter-quartile value (Q1)
$\mathbf{T_8}$	Variance	$\mathbf{T_{20}}$	Second inter-quartile value (Q2)
$\mathbf{T_9}$	Root mean square	$\mathbf{T_{21}}$	Third inter-quartile value (Q3)
$\mathbf{T_{10}}$	Waveform length	$\mathbf{T_{22}}$	Mode value
$\mathbf{T_{11}}$	Average amplitude change value	$\mathbf{T_{23}}$	Zero-crossing value
$\mathbf{T_{12}}$	Absolute difference in standart deviation	$\mathbf{T_{24}}$	Slope-change value

In another feature set, frequency-domain feature set, the t-test selected features were investigated. The list of 15 frequency-domain features with their abbreviations are given in Table 3.11. Channel-based t-test selected statistically significant frequency-domain feature distribution is given in Table 3.12. A total of 28 frequency-domain features are indicated as statistically significant features with the application of t-test.

Table 3.10: Channel-based t-test selected statistically significant feature distribution for binary extremity movement classification in time-domain feature set.

F_{id}	Channels																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	T
T_1																							0
T_2																							3
T_3																							12
T_4																							0
T_5																							0
T_6																							0
T_7																							0
T_8																							0
T_9																							0
T_{10}																							0
T_{11}																							0
T_{12}																							0
T_{13}																							4
T_{14}																							3
T_{15}																							0
T_{16}																							0
T_{17}																							0
T_{18}																							0
T_{19}																							5
T_{20}																							13
T_{21}																							1
T_{22}																							0
T_{23}																							0
T_{24}																							3
T	0	0	0	2	4	5	0	0	0	0	3	5	5	0	1	2	3	5	2	2	2	3	44

Table 3.11: FFT-based frequency and WT-based time-frequency domain features.

Frequency-domain and time-frequency domain features			
F_1, W_1	Energy of delta band	**F_9, W_9**	Entropy of alpha band
F_2, W_2	Variance of delta band	**F_{10}, W_{10}**	Energy of beta band
F_3, W_3	Entropy of delta band	**F_{11}, W_{11}**	Variance of beta band
F_4, W_4	Energy of theta band	**F_{12}, W_{12}**	Entropy of beta band
F_5, W_5	Variance of theta band	**F_{13}, W_{13}**	Energy of gamma band
F_6, W_6	Entropy of theta band	**F_{14}, W_{14}**	Variance of gamma band
F_7, W_7	Energy of alpha band	**F_{15}, W_{15}**	Entropy of gamma band
F_8, W_8	Variance of alpha band		

Table 3.12: Channel-based t-test selected statistically significant feature distribution for binary extremity movement classification in frequency-domain feature set.

F_{id}	Channels																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	T
F_1																							0
F_2																							0
F_3																							0
F_4																							0
F_5																							0
F_6																							5
F_7																							8
F_8																							5
F_9																							7
F_{10}																							0
F_{11}																							2
F_{12}																							0
F_{13}																							0
F_{14}																							0
F_{15}																							1
T	1	1	0	0	2	1	1	1	0	0	1	4	3	1	0	0	2	5	1	1	2	1	28

Among the 15 different time-domain features, the selected features are entropy of theta band (in 5 EEG channels), energy of alpha band (in 8 EEG channels), variance of alpha band (in 5 EEG channels), entropy of alpha band (in 7 EEG channels), variance of beta band (in 2 EEG channels), and entropy of gamma band (in an EEG channel). In literature it has been noted that alpha and beta rhytsms may be related motor activities [1]. The alpha rhythms reflect visual processing and can be also associated with memory brain function. Also, Mu rhythms may be available in the same frequency-range as alpha rhythms. Mu rhythms are strongly related to motor activities and, in some conditions, appear to correlate with beta rhythms. Beta rhythms are strongly related to motor activities. In our study, supporting the literature, statistically significant features were determined as frequency-domain features obtained using the alpha band. Therefore, it has been observed that classification performance improves in most classifiers by selecting effective frequency band-based features. On the other hand, among 22 EEG channels, more statistically significant features were selected from some channels (12th, 13th, and 18th EEG channels). In studies in the literature, certain channels (8th, 10th, and 12th) were identified as effective channels and they were used for extremity movement classification [128]. Selecting more statistically significant features from certain channels such as 12th EEG channels may also have improved classifier performance.

Table 3.13: Channel-based t-test selected statistically significant feature distribution for binary extremity movement classification in time-frequency domain feature set.

F_{id}	Channels																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	T
W_1																							0
W_2																							0
W_3																							3
W_4																							0
W_5																							0
W_6																							2
W_7																							1
W_8																							0
W_9																							0
W_{10}																							0
W_{11}																							0
W_{12}																							2
W_{13}																							0
W_{14}																							0
W_{15}																							5
T	0	1	0	2	0	0	1	0	0	1	0	1	0	1	1	0	0	1	0	1	2	1	13

Then, the t-test selected features were investigated in WT-based time-frequency domain feature set. The list of 15 time-frequency domain features with their abbreviations are given above with Table 3.11. Channel-based t-test selected statistically significant time-frequency domain feature distribution is given in Table 3.13. A total of 13 time-frequency domain features are indicated as statistically significant features with the application of t-test. Among the 15 different time-fequency domain features, the selected features are entropy of delta band (in 3 EEG channels), entropy of theta band (in 2 EEG channels), energy of alpha band (in an EEG channel), entropy of beta band (in 2 EEG channels), and entropy of gamma band (in 5 EEG channels). Contrary to the literature [1], it was observed that the t-test improved the classifier performance by not selecting the features obtained from the alpha and beta bands which are related to motor activities, but only by making a selection based on the entropy features of the other bands. On the other hand, among 22 EEG channels, more statistically significant features were selected from some channels (4^{th} and 21^{th} EEG channels). Contrary to the literature [128], it was observed that the t-test improved the classifier performance by not selecting the features obtained from certain channels which are indicated as effective channels, but only by making a selection based on the features of 4^{th} and 21^{th} EEG channels.

Table 3.14: Poincare plot-based non-linear features.

Non-linear features			
P_1	SD_1 where lag=1	**P_5**	SD_1 where lag=9
P_2	SD_2 where lag=1	**P_6**	$SD_2$where lag=9
P_3	SD_1SD_2 where lag=1	**P_7**	SD_1SD_2 where lag=9
P_4	SD_1/SD_2 where lag=1	**P_8**	SD_1/SD_2 where lag=9

Finally, the selected statistically significant non-linear features were investigated in the second combination feature set (TD+FD+WT+P). The list of 8 non-linear features with their abbreviations are given in Table 3.14. Channel-based t-test selected statistically significant non-linear feature distribution is given in Table 3.15. A total of 6 non-linear features are indicated as statistically significant features with the application of t-test. Among the 8 different non-linear features, the selected features are SD_1 where lag=9 (in an EEG channel) and SD_1/SD_2 where lag=9 (in 5 EEG channels). It has been observed that among the non-linear features obtained for both lag values, the non-linear features obtained for lag = 9 were selected only as statistically significant features. As a result, this explains why we add the features obtained for the lag = 9 case to our second combination set. However, including selected non-linear features to the combination set generally improved classifier performance. On the other hand, among 22 EEG channels, statistically significant features were selected from a channel (12th EEG channel). This channel indicated as effective channel for MI task classification in the literature [128]. Therefore, selecting statistically significant features from it may also have improved classifier performance.

Table 3.15: Channel-based t-test selected statistically significant feature distribution for binary extremity movement classification in non-linear feature set.

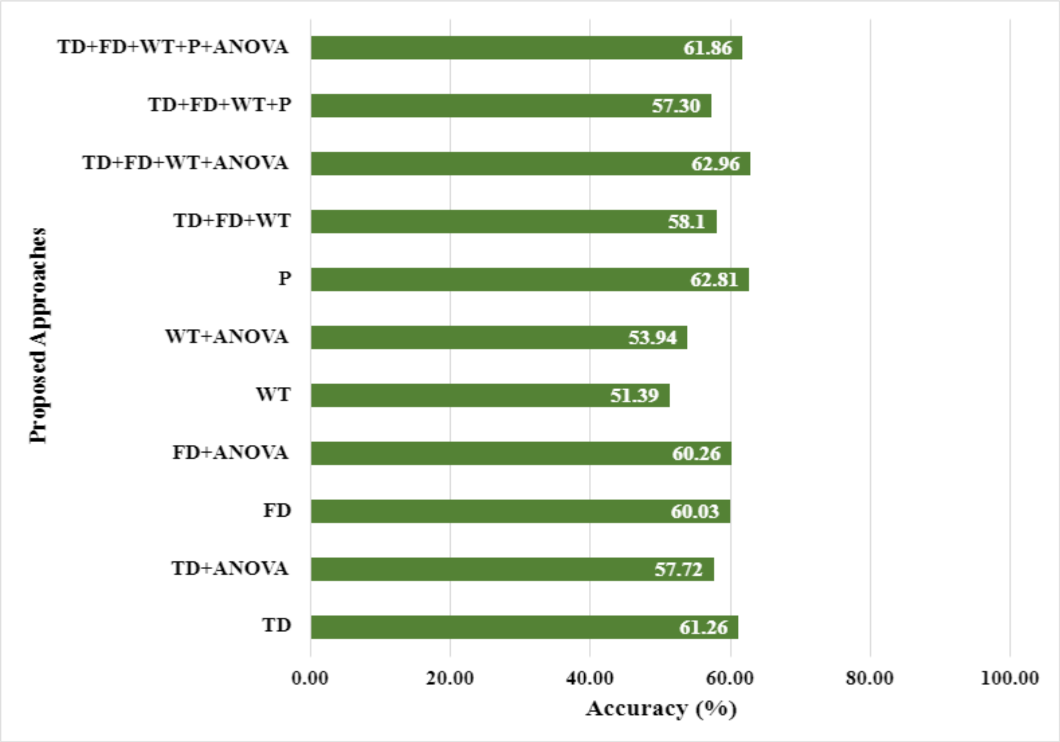
F_{id}	Channels																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	T
P_1																							0
P_2																							0
P_3																							0
P_4																							0
P_5																							1
P_6																							0
P_7																							0
P_8																							5
T	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	2	0	0	0	0	6

In our proposed binary-class extremity movement task classification studies, our main purpose is to introduce different feature extraction-based approaches and investigate the effects of these feature sets and the statistically significance-based feature selection on the classification performance. In our experiments we included the Poincare plot-based non-linear features that have not been used in MI task classification in previous studies.

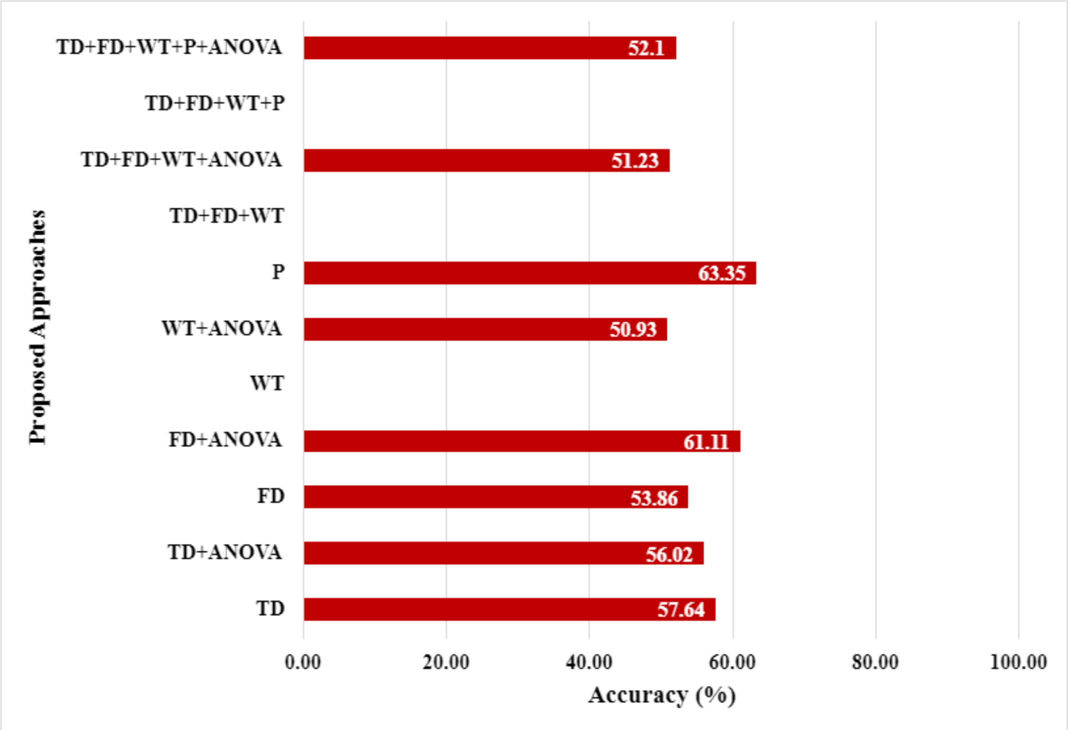
We performed the proposed four different feature extraction approaches on classification of right and left hands MI task 22-channel EEG signals supplied from open-available dataset. 24 time-domain, 15 frequency-domain, 15 time-frequency domain and 4 non-linear features are extracted from each EEG segments. These time-domain, frequency-domain, time-frequency domain, non-linear feature sets their two different combination feature sets were classified utilizing DT, DA, NB, SVM, LR, k-NN, EL, NN, and KA, and performances of different feature sets were compared. Additionally, the independent t-test was applied to select features in the proposed feature sets and the effectiveness of this method is analyzed in all feature sets with same classifiers.

Among all feature sets, performance of time-frequency feature set was observed to be poor for binary-class extremity movement task classification and the performance of non-linear feature sets was found to be higher especially for all classifiers except NB. The most successful non-linear feature set is 6^{th} feature set including non-linear features for lag=6 condition. The highest accuracy value of binary classification is provided by using this non-linear feature set. Therefore, the successful non-linear feature sets revealed that MI tasks lead to distinctive and effective differences in the non-linear dynamics embedded in EEG signals. When the analyzes performed with two different combination sets were compared, it was observed that, contrary to the individual success of the non-linear feature sets, better performances were achieved with the 1st combination feature set (TD+FD+WT), in which Poincare measurements were not included.

In addition, when the effects of the independent t-test were evaluated, we noted that performance of this feature selection with time-domain feature set was observed to be poor and sufficient improvements in classifier performances have not been achieved. On contrary to the classification performed using statistically significant features from time-domain feature set, the proposed the independent t-test based feature selection generally improves classifier performance in other feature sets-based binary classifications. The maximum accuracy value in binary classification is evaluated with DA algorithm for non-linear feature set, but the highest accuracy value in different feature sets is generally evaluated using EL algorithm as shown in Figure 3.3a and Figure 3.3b. The detailed comparision of accuracy values of proposed approaches using EL and DA algorithms are given in Figure 3.3a and Figure 3.3b, respectively.



(a) EL-based binary extremity movement classification results.



(b) DA-based binary extremity movement classification results.

Figure 3.3: Comparing of accuracy values of proposed binary-class extremity movement task classification approaches using (a) EL algorithm and (b) DA algorithm.

Table 3.16: Performance comparison of binary-class extremity movement task classification studies.

Ref.	Subject condition	Number of channels	Number of classes	Proposed methods	Classifier	Accuracy (%)
[38]	SI/9	2	2	WPD	Random forest	68.32
[81]	SI/9	22	2	Time-domain parameters FFT T-test	EL	62.52
[128]	SI/9	3	2	STFT	CNN	74.20
[129]	SI/9	22	2	WT	CNN	69.00
This study	**SI/9**	**22**	**2**	**TD**	**EL**	**62.26**
				FD+T-test	**LR**	**61.34**
				WT+ T-test	**NB**	**54.75**
				P	**LDA**	**63.35**
				TD+FD+WT+ T-test	**EL**	**62.96**
				TD+FD+WT+P+ T-test	**EL**	**61.86**

In Table 3.16, some of the previous binary-class extremity movement classification studies are summarized and their performances are compared with that of the proposed study. In [38], binary-class extremity movement classification was presented using WPD. They selected only two EEG channels before feature extraction step and Random forest-based classification results reported as 68.32%. However, they used certain channels eliminating information of other channels. When we examine the channel-based distribution of statistically significant features in our studies, we observed that significant features are selected from different features or different channels in different data sets and that the same EEG channels are not selected in all feature sets. Therefore, when we select certain channels and work with them as that of study, we would not have captured the significant features in some channels. In [128], STFT-based binary-class MI task classification was performed using 3 EEG channels of data set. Higher accuracy value was yielded than our presented studies. However, in their study channel selection step was performed which is not the stage in our studies. We used 22 EEG channels with high and low classification performance, which slightly decreases the overall motor imagery task classification performance, but eliminates a channel selection phase. In addition to channel reduction, their proposed study includes high complexity in terms of CNN-based feature extraction and classification due to fact that its deep neural layer structures increase training time. In another CNN-based classification study [129], WT algorithm was used for feature extraction. Although it has high computational complexity, the classification result was reported as 69.00%. In our previous study [81], we investigated the statistically significant time-domain and frequency-domain features in binary-class extremity movement classification using EL algorithm. We observed that t-test improved classifier performance. In addition to these time-domain and frequency-domain feature extraction approaches, different feature extraction methods, which clearly have the computational advantages, were investigated by us. Thus, the above encouraging experimental results together with the computational advantages, indicate that the proposed Poincare plot measures and the combination of different feature extraction approaches may be used to analysis of non-stationary EEG signals.

3.5 Results and Discussions of Multi-Class Extremity Movement Task Classification

In the multi-class extremity movement task classification, we aim at discriminating the four different MI task EEG segments utilizing different feature extraction-based approaches. Four different (right hand, left hand, both feet, and tongue) extremity movement MI tasks segments of 22-channel EEG recordings obtained from 9 subjects are used to extract features. Time-domain, frequency-domain, time-frequency domain and non-linear feature sets are obtained from these EEG segments. Time-domain (24 different statistical and amplitude-based measures), frequency-domain (energy, variance, and entropy measures of FT-based five different EEG sub-bands), time-frequency domain (energy, variance, and entropy measures of WT-based five different EEG sub-bands), and non-linear (4 different Poincare plot measures) feature sets were created from 22-channel EEG signals. The effect of these four different feature sets and their two different combination sets are investigated as previous binary classification studies. In addition, the effectiveness of ANOVA-based feature selection process is investigated with all extracted feature sets. Finally, 8 different machine learning algorithms are DT, DA, NB, SVM, k-NN, EL, NN, and KA were applied to classify these feature sets and the results of each of them were analyzed. All signal processing (signal segmentation, feature extraction, feature selection, and classification) and performance analyzes were implemented using MATLAB software.

Performance evaluation results of our proposed approach including different feature sets, ANOVA-based feature selection, and various classifiers are given in Tables 3.17-3.22. In these tables TD, FD, WT, and P denote that the features for classifications using the time-domain, frequency-domain, time-frequency domain, and non-linear information, respectively. On the other hand, TD+ANOVA, FD+ANOVA, and WT+ANOVA indicate that the ANOVA selected statistically significant features for classifications using the time-domain, frequency-domain, and time-frequency domain, respectively. The classifications performed using the first combination feature set including time-domain, frequency-domain, and time-frequency domain features are denoted as TD+FD+WT and the classifications performed using the ANOVA-selected statistically significant features of this combination set are denoted as TD+FD+WT+ANOVA. The classifications performed using the second combination feature set including time-domain, frequency-domain, time-frequency domain, and non-linear features are denoted as TD+FD+WT+P and the classifications performed using the ANOVA selected statistically significant features of this combination set are denoted as TD+FD+WT+P+ANOVA. The boldface characters in table cells denote the best classification performance for each approach and classification algorithms (in Tables 3.17-3.22).

The performance evaluation of all time-domain features and the ANOVA-selected statistically significant time-domain features for four-class MI task classification is summarized in Table 3.17. We obtained the highest accuracy value of 44.38% using all time-domain features evaluated from EEG segments and EL algorithm while the NB algorithm performed the worst accuracy value of 29.40% for the same features. When the ANOVA-selected statistically significant time-domain features evaluated using various classifiers, we achieved the highest 43.91% classification accuracy using EL algorithm and the worst accuracy value of 29.40% using NB algorithm. To investigate the effectiveness of the independent t-test selection process, we analyzed and compared performance results of TD and TD+ANOVA classifications. It was observed that ANOVA based feature selection process improved the performance in 5 classifiers, decreased the performance in 2 classifiers, and did not change the performance in 1 classifier. Results of all classification using time-domain and statistically significant time-domain features are given in Table 3.17.

Table 3.17: Performance results (%) for multi-class extremity movement task classification using time-domain feature set.

Models	**Accuracy**	
	TD	**TD+ANOVA**
Decision Tree	31.00	31.10
Discriminant Analysis	41.90	44.00
Naive Bayes	29.40	29.40
Support Vector Machine	40.28	43.12
k-Nearest Neighbours	32.30	33.40
Ensemble Learning	**44.38**	**43.91**
Neural Networks	39.89	40.86
Kernel Approximation	32.48	31.87

The performance evaluation of all frequency-domain features and the ANOVA-selected statistically significant frequency-domain features for four-class MI task classification is summarized in Table 3.18. We obtained the highest accuracy value of 35.76% using all frequency-domain features evaluated from EEG segments and EL algorithm while the NB algorithm performed the worst accuracy value of 28.59% for the same features. When the ANOVA-selected statistically significant frequency-domain features evaluated using various classifiers, we achieved the highest 38.46% classification accuracy using EL algorithm and the worst accuracy value of 29.09% using NB algorithm. To discover the effectiveness of the ANOVA-based feature selection process in frequency-domain feature set, we analyzed and compared performance results of FD and FD+ANOVA classifications. It was observed that ANOVA-based feature selection improved classifier performances in all classifiers except a classifier (KA). Results of all classification using frequency-domain and statistically significant frequency-domain features are given in Table 3.18.

Table 3.18: Performance results (%) for multi-class extremity movement task classification using frequency-domain feature set.

Models	Accuracy	
	FD	FD+ANOVA
Decision Tree	31.40	31.44
Discriminant Analysis	34.41	37.89
Naive Bayes	28.59	29.09
Support Vector Machine	33.14	37.69
k-Nearest Neighbours	29.28	29.98
Ensemble Learning	**35.76**	**38.46**
Neural Networks	33.68	36.38
Kernel Approximation	32.18	30.94

The performance evaluation of all time-frequency domain features and the ANOVA-selected statistically significant time-frequency-domain features for four-class MI task classification is summarized in Table 3.19. We obtained the highest accuracy value of 28.63% using all time-frequency domain features evaluated from EEG segments and EL algorithm while the SVM, k-NN, and KA algorithms performed the worst accuracy value of 24.81% for the same features. In DA algorithm, we did not perform classification using all time-frequency domain features, because the proposed feature set is not suitable for the classifier structure. When the ANOVA-selected statistically significant time-frequency domain features evaluated using various classifiers, we achieved the highest 34.34% classification accuracy using EL algorithm and the worst accuracy value of 25.42% using DA algorithm. To discover the effectiveness of the ANOVA-based feature selection process in time-frequency domain feature set, we analyzed and compared performance results of WT and WT+ANOVA classifications. When the performed classifications were examined for the case where all features were used, it was observed that the ANOVA-based feature selection method increased the performance in all classifiers except one out of 7 classifiers. Results of all classification using time-frequency domain and statistically significant time-frequency domain features are given in Table 3.19.

Table 3.19: Performance results (%) for multi-class extremity movement task classification using the time-frequency domain feature set.

Models	Accuracy	
	WT	**WT+ANOVA**
Decision Tree	28.32	28.74
Discriminant Analysis	N/A	25.42
Naive Bayes	28.20	28.16
Support Vector Machine	24.81	25.73
k-Nearest Neighbours	24.81	25.54
Ensemble Learning	**28.63**	**34.34**
Neural Networks	25.00	25.62
Kernel Approximation	24.81	25.81

The performance evaluation of all non-linear feature sets with various classifiers is summarized in Table 3.20. In this table, Lag (1)-Lag (10) indicate that the features for classifications are evaluated by using the corresponding lag value. Additionally, "All lags" indicates that the classifications are carried out using the combination feature set provided by combining the features from 10 different lag values. The results revealed that the non-linear feature set extracted for All lags condition achieved the highest accuracy value of 47.08% using SVM classifier and the worst accuracy value of 26.80% is evaluated using lag=1 feature set and NB classifier. On the other hand, we analyzed the effectiveness of different lag values for 8 classifiers. We observed that the higher classification performance is obtained in 1 (DA) classifiers using feature set for lag=7 condition, 1 (k-NN) classifiers using feature set for lag=8 condition, in 2 (DA, and NB) classifiers using feature set for lag=9 condition, 1 (KA) classifiers using feature set for lag=10 condition, and in 4 (DT, SVM, EL, and NN) classifiers using feature set for All lags condition. Among all non-linear feature sets, the most successful feature set is evaluated as the All lags feature set generated combining of all non-linear feature sets. When we analyzed the other 10 non-linear sets without including the combination feature set, it was seen that the highest results were achieved with the 9th feature set generated where lag=9.

Table 3.20: Performance results (%) for multi-class extremity movement task classification using the non-linear feature set.

Model	Lag (1)	Lag (2)	Lag (3)	Lag (4)	Lag (5)	Lag (6)	Lag (7)	Lag (8)	Lag (9)	Lag (10)	All lags
Decision Tree	29.10	30.10	29.60	29.30	29.40	30.60	30.50	31.40	31.50	31.10	**31.90**
Discriminant Analysis	40.00	40.50	40.00	41.70	40.10	41.90	**42.70**	42.30	**42.70**	41.70	40.20
Naive Bayes	26.80	28.30	27.90	28.60	28.00	29.00	28.00	28.20	**30.00**	28.40	28.30
Support Vector Machine	41.01	43.16	44.30	44.51	42.98	43.07	44.36	44.48	43.41	43.78	**47.08**
k-Nearest Neighbours	32.20	32.10	32.50	32.40	32.10	32.40	32.20	**33.30**	33.10	33.00	32.30
Ensemble Learning	39.59	41.20	41.53	42.08	40.19	42.17	42.84	42.94	43.42	42.27	**46.06**
Neural Networks	39.97	41.71	41.59	42.05	41.00	41.20	42.00	41.47	40.35	41.05	**45.18**
Kernel Approximation	30.13	31.52	31.40	31.79	32.29	32.25	31.48	31.60	30.98	**32.64**	30.63

Table 3.21: Performance results (%) for multi-class extremity movement task classification using the combined (TD+FD+WT) feature set.

Models	Accuracy	
	TD+FD+WT	**TD+FD+WT+ANOVA**
Decision Tree	34.38	34.34
Discriminant Analysis	N/A	25.96
Naive Bayes	29.09	29.51
Support Vector Machine	24.81	26.54
k-Nearest Neighbours	24.81	25.62
Ensemble Learning	**35.73**	**44.33**
Neural Networks	25.00	26.54
Kernel Approximation	24.81	25.46

In order to compare the effectiveness of different feature sets, the classification is carried out with the combination of time-domain, frequency-domain, and time-frequency domain features. The performance evaluation results of this combination feature set and the ANOVA-selected feature set from the combined feature set are summarized in Table 3.21. In classification performed using combined feature set, the EL algorithm obtained the maximum accuracy (35.73%) and SVM, k-NN, and KA algorithms obtained the worst accuracy (24.81%) using same features. In DA algorithm, we did not perform classification using all combination set feature, because the proposed feature set is not suitable for the classifier structure. On the other hand, in the classifications performed using the ANOVA-selected feature set, the EL algorithm provided the maximum accuracy (44.33%) and KA provided the worst accuracy (25.46%) using same features. To discover the effectiveness of the ANOVA-based feature selection process in the combination feature set, we analyzed and compared performance results of TD+FD+WT and TD+FD+WT+ANOVA classifications. When the performed classifications were examined for the case where all features were used, it was observed that the ANOVA-based feature selection method increased the performance in all classifiers except one out of 7 classifiers. Results of all classification using the combination set and the selected statistically significant combination set features are given in Table 3.21.

Table 3.22: Performance results (%) for multi-class extremity movement task classification using the combined (TD+FD+WT+P) feature set.

	Accuracy	
Models	**TD+FD+WT+P**	**TD+FD+WT+P+ANOVA**
Decision Tree	34.50	34.50
Discriminant Analysis	N/A	27.31
Naive Bayes	27.90	29.43
Support Vector Machine	25.00	29.43
k-Nearest Neighbours	24.90	26.21
Ensemble Learning	**35.60**	**47.36**
Neural Networks	24.90	27.55
Kernel Approximation	24.90	25.89

The classification is performed in our second combination set including time-domain, frequency-domain, time-frequency domain, and non-linear features. We added non-linear feature sets extracted for lag=1 and lag=9 conditions, into our previous combination set. The performance evaluation results of this combination feature set and the ANOVA-selected feature set from this combined feature set are summarized in Table 3.22. In classification performed using combined feature set, the EL algorithm obtained the maximum accuracy (35.60%) and k-NN, NN, and KA obtained the worst accuracy (24.90%) using same features. On the other hand, in the classifications performed using the ANOVA-selected feature set, the EL algorithm provided the maximum accuracy (47.36%) and KA provided the worst accuracy (25.89%) using same features. To discover the effectiveness of the ANOVA-based feature selection process in the combination feature set, we analyzed and compared performance results of TD+FD+WT+P and TD+FD+WT+P+ANOVA classifications. When the performed classifications were examined for the case where all features were used, it was observed that the ANOVA-based feature selection method increased the performance in all classifiers except one out of 7 classifiers. Results of all classification using our second combination set and the selected statistically significant combination set features are given in Table 3.22.

ANOVA-selected features in four different sets are analyzed to investigate effects of different features and EEG channels for multi-class extremity movement classification. Firstly, we investigated the ANOVA-selected time-domain features. The list of 24 different time-domain features with their abbreviations are available above with Table 3.9. Channel-based ANOVA-selected statistically significant time-domain feature distribution is given in Table 3.23. A total of 345 time-domain features were indicated as statistically-significant features using ANOVA. Among 24 different time-domain features, some features such as minumu value, maximum value, mean, standard deviation value, integrated EEG value, mean absolute value, simple square integral, variance, root mean square, skewness, Hjorth parameters, signal range, Q1, Q2, zero crossing value, and slope-change value were mostly selected as statistically significant features from almost all channels. As can be clearly observed in the Table 3.23, some features were not selected as statistically significant features in any channel. When the effectiveness of 22 EEG channels was examined, it was observed that statistically significant features were selected intensively from all channels, and there was no density in certain channels. As a result, it has been observed that classifier performance has generally improved by determining statistically significant features from all channels and certain features with ANOVA.

Table 3.23: Channel-based ANOVA-selected statistically significant feature distribution for multi-class extremity movement task classification in time-domain feature set.

F_{id}	Channels																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	T
T_1	■	■	■	■				■	■	■	■	■		■	■	■	■	■	■	■	■	■	18
T_2		■					■	■			■	■	■	■	■	■	■	■	■	■	■	■	15
T_3	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	22
T_4	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	22
T_5	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	22
T_6	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	22
T_7	■	■	■	■	■		■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	20
T_8	■	■	■	■	■		■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	20
T_9	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	22
T_{10}																							0
T_{11}																							0
T_{12}																							0
T_{13}																							0
T_{14}														■				■	■	■			4
T_{15}	■	■	■	■	■		■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	20
T_{16}							■	■						■	■	■	■	■	■	■	■	■	11
T_{17}	■	■					■	■						■	■	■			■	■	■	■	11
T_{18}		■	■				■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	17
T_{19}	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	22
T_{20}	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	22
T_{21}	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	22
T_{22}		■																					1
T_{23}							■	■						■	■	■		■	■	■	■	■	10
T_{24}	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	22
T	14	16	14	13	11	9	17	18	14	14	15	15	10	19	18	18	16	18	19	19	18	18	345

Channel-based ANOVA-selected statistically significant frequency-domain feature distribution for multi-class extremity movement task classification is given in Table 3.24. A total of 102 frequency-domain features were indicated as statistically-significant features using ANOVA. Among 15 different frequency frequency-domain features, the mostly selected statistically significant features were the energy and variance values of theta, alpha, and beta bands. These features are selected in too many channels. On the other hand, among 22 EEG channels, more statistically significant features were selected from some channels (10^{th}, 16^{th}, 18^{th}, 20^{th}, 21^{th}, and 22^{th} EEG channels). In studies in the literature, 10^{th} EEG channel was identified as effective channels and they were used for MI task classification [128]. Selecting more statistically significant features from certain channels such as 10^{th} EEG channels and certain EEG subbands such as alpha and beta bands which are associated with motor activities may have improved classifier performance.

Table 3.24: Channel-based ANOVA-selected statistically significant feature distribution for multi-class extremity movement task classification in frequency-domain feature set.

F_{id}	Channels																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	T
F_1																							0
F_2																							0
F_3																							0
F_4																							10
F_5																							5
F_6																							3
F_7																							22
F_8																							22
F_9																							4
F_{10}																							18
F_{11}																							17
F_{12}																							0
F_{13}																							0
F_{14}																							0
F_{15}																							1
T	2	4	4	4	4	2	3	4	5	6	5	5	3	4	5	6	5	6	5	6	6	8	102

Channel-based ANOVA-selected statistically significant WT-based time-frequency domain feature distribution for multi-class extremity movement task classification is given in Table 3.25. A total of 104 time-frequency domain features were indicated as statistically-significant features using ANOVA. Among 15 different frequency frequency-domain features, the mostly selected statistically significant features were the energy and variance values of delta, theta, and alpha bands. These features are selected in too many channels. On the other hand, among 22 EEG channels, more statistically significant features were selected from some channels (14^{th}, 19^{th}, 20^{th}, 21^{th}, and 22^{th} EEG channels). However, it was observed that statistically significant time-frequency features were generally selected from all channels. Selecting statistically significant features from all EEG channels and certain EEG subbands such as alpha band which is associated with motor activities may have improved classifier performance.

Table 3.25: Channel-based t-test selected statistically significant feature distribution for multi-class extremity movement task classification in time-frequency domain feature set.

F_{id}	Channels																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	T
W_1																							5
W_2																							5
W_3																							1
W_4																							22
W_5																							22
W_6																							0
W_7																							22
W_8																							22
W_9																							1
W_{10}																							0
W_{11}																							0
W_{12}																							0
W_{13}																							0
W_{14}																							0
W_{15}																							4
T	4	5	4	4	4	4	5	4	4	4	4	5	4	6	4	4	4	4	7	7	6	7	104

Table 3.26: Channel-based t-test selected statistically significant feature distribution for multi-class extremity movement task classification in non-linear feature set.

F_{id}	Channels																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	T
P_1																							5
P_2																							7
P_3																							8
P_4																							6
P_5																							10
P_6																							10
P_7																							7
P_8																							8
T	5	2	0	2	5	1	1	5	5	1	3	6	3	0	4	5	2	0	5	5	1	0	61

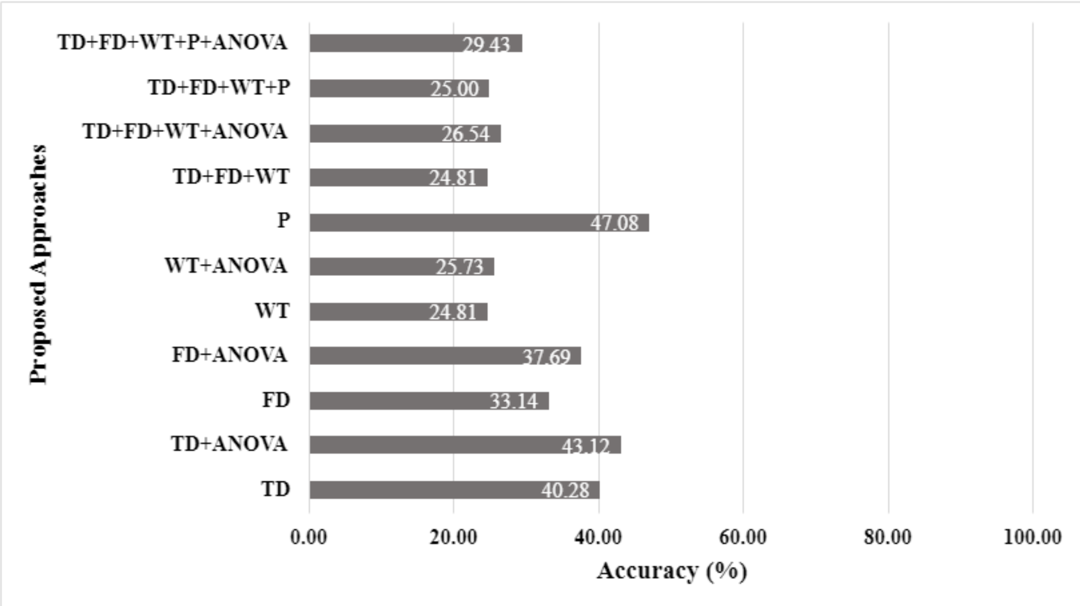
Finally, the selected statistically significant non-linear features were investigated in the second combination feature set (TD+FD+WT+P). The list of 8 non-linear features with their abbreviations are given above with Table 3.14. Channel-based ANOVA-selected statistically significant non-linear feature distribution is given in Table 3.14. A total of 61 non-linear features are indicated as statistically significant features with the application of ANOVA test. When the selected statistically significant non-linear features were examined, it was observed that balanced selections were made from all features, not specific features. On the other hand, among 22 EEG channels, statistically significant features were selected from almost all of the EEG channels. However, from some channels such as 8^{th} and 12^{th} EEG channels, more statistically significant were selected. These channels were indicated as effective EEG channels in literature [128]. Therofore, lots of selections on these channels and there is a balanced distribution of feature selection from all feature may have improved classifier performance in our second combination feature set (TD+FD+WT+P).

In our proposed four-class extremity movements task classification studies, we investigated the effects of various feature sets and statistically significance-based feature selection method on the classification performance. In addition to the classical feature extraction approaches which are time-domain, frequency-domain, and time-frequency domain-based evaluations, we investigate the effects of non-linear features (Poincare plot-based measures) for four-class MI task classification. Additionally, we combined different feature sets and applied ANOVA-based feature selection process to improve classifier performance determining effective and relevant features from EEG signals.

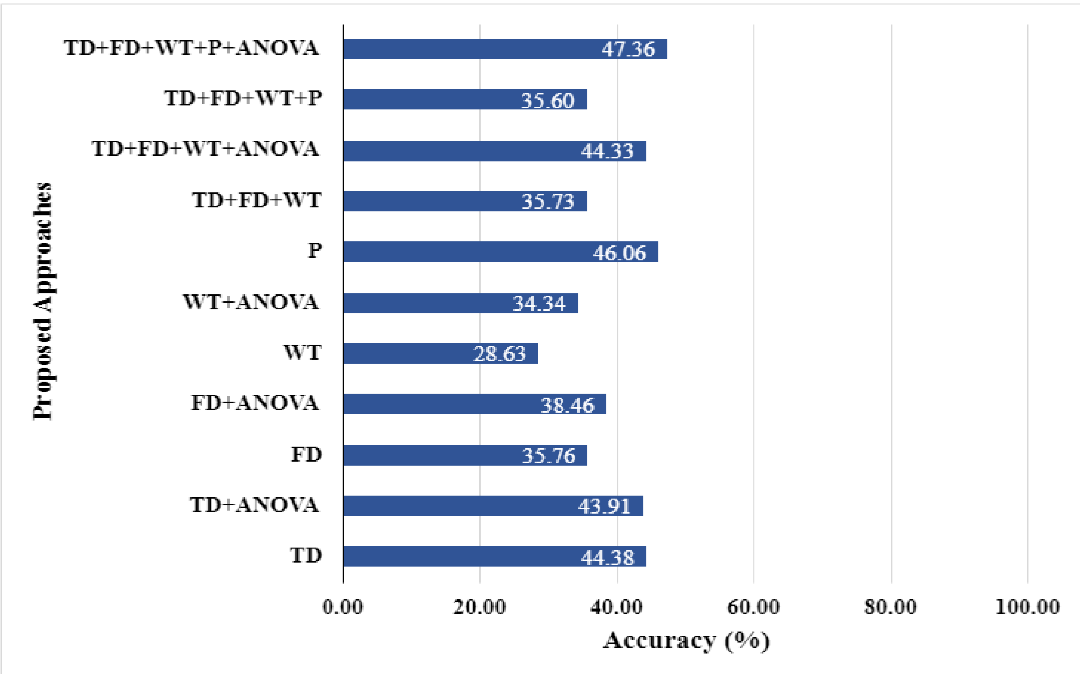
We applied the proposed four different feature extraction approaches for classification of right hand, left hand, both feet, and tongue MI tasks of 22-channel EEG signals after obtaining of 3 sec MI EEG segments. A total of 24 time-domain, 15 frequency-domain, 15 time-frequency domain and 4 non-linear features are evluated from each EEG segment. These time-domain, frequency-domain, time-frequency domain, and non-linear feature sets, their two different combination feature sets, and ANOVA-selected statistically significant feature sets of all feature sets were classified utilizing DT, DA, NB, SVM, k-NN, EL, NN, and KA, and performances of different feature sets and the effectiveness of ANOVA were investigated and compared.

Among all feature sets, performance of time-frequency feature set was observed to be poor for four-class extremity movement task classification the performance of non-linear feature sets was found to be higher especially for SVM and EL classifiers. When we examined the 4 feature sets apart from the combinations, we observed that the most successful is non-linear feature set which is defined as All lags feature set including non-linear features for all lag condition. The highest accuracy value of multi-class classification is provided by using this non-linear feature set and SVM algorithm. Therefore, the successful non-linear feature sets revealed that MI tasks lead to distinctive and effective differences in the non-linear dynamics embedded in EEG signals. The performance evaluation of all proposed approaches using SVM algorithm is given in Figure 3.3a.

When the analyzes performed with two different combination sets were compared, it was observed that, better performances were achieved with the 2nd combination feature set (TD+FD+WT+P), in which Poincare measurements were included. At the same time, the highest performance value of the four-class MI task classification studies is achieved with this combination feature set and EL. It has been observed that, despite the high classification performance of the SVM algorithm in the non-linear data group among the four feature sets, it does not show the same performance in the combination feature set. In addition, it has been noticed that the highest performance in the proposed approaches was generally achieved with the EL algorithm in the classifications carried out on all features sets (time-domain, frequency-domain, time-frequency domain, and combination sets) except the non-linear feature sets.



(a)



(b)

Figure 3.4: Comparing of accuracy values of proposed multi-class extremity movement task classification approaches using (a) SVM algorithm and (b) EL algorithm.

The detailed comparision of accuracy values of proposed approaches using SVM and EL algorithms are given in Figure 3.4a and Figure 3.4b, respectively. In the given classifier performances, it is clearly observed with which feature set the highest classification performance is achieved.

In addition, when the effects of the ANOVA-based feature selection were investigated, we noted that performance of this feature selection with all feature sets especially combination feature sets is improved the four-class MI task classification performance determining the effective and relevant features. The maximum accuracy value in multi-class classification is evaluated with EL algorithm and TD+FD+WT+P+ANOVA combination feature set as given in Figure 3.4b. Therefore, the combination of different feature extraction methods and the statistically significance-based feature selection method can be improved classifier performance and diminished classifier complexity selecting small number of discriminative features.

Table 3.27: Performance comparison of multi-class extremity movement task classification studies.

Ref.	Subject condition	Number of channels	Number of classes	Proposed methods	Classifier	Accuracy (%)
[130]	SI/9	22	4	FBCSP Energy-based features	CNN	70.60
[131]	SI/9	8	4	FFT Channel variance features PCA	SVM	56.00
[132]	SI/9	22	4	CSP	Fuzzy logic system	65.00
[133]	SI/9	26	4	CSP Band power	LDA	51.67
[134]	SI/10	64	4	Time domain parameters	LDA	58.30
This study	**SI/9**	**22**	**4**	**TD**	**EL**	**44.38**
				FD+ANOVA	**EL**	**38.46**
				WT+ANOVA	**EL**	**34.34**
				P	**SVM**	**47.08**
				TD+FD+WT+ANOVA	**EL**	**44.33**
				TD+FD+WT+P+ANOVA	**EL**	**47.36**

Performance comparison of multi-class extremity movement task classification studies conducted in the literature is demonstrated in Table 3.27. In [130], a CNN-based approach is presented. FBCSP and energy-based features of EEG signals were used as input. The accuracy of 70.60% were achieved for subject-independent analysis of 9 subjects. In that study, higher MI task classification accuracy was achieved compared to our presented studies at the expense of computationally expensive feature extraction and classification step. In another study [131], channel variance-based feature extraction and PCA-based feature selection were used. The extracted features were classified with accuracy of 56.00% using SVM algorithm. Higher accuracy value was obtained in that study than that of our study. While EEG data of 8 channels are investigated in that study, EEG data of 22 channels are examined in our analysis. In another studies [132-134] CSP, band power and CSP, time-domain parameters have been used for feature extraction for each study respectively. In these studies, higher accuracy values were avhieved compared to our studies. However, performance values are not very high (over 70.00%). Our multi-class extremity movement classification study is different from studies that use specific and same feature extraction methods in that it works with very different feature sets. This study is the first to include the Poincare plot measures-based non-linear feature set in the feature sets examined and investigate its effectiveness alone and in different combinations. Promising results were obtained with the different proposed feature sets and the ANOVA-based feature selection method used. Additionally, detailed research was conducted on the effectiveness of channels and different features by examining the ANOVA-based selected statitistically significant features.

Chapter 4

Classification of Finger Movement

In this section of the thesis, four different feature extraction approaches and combinations of different approaches, and two different feature extraction approaches known as statistically significance-based feature selection and PCA are applied to classify EEG segments of finger movement. In addition to these feature extraction approaches, ITD-based feature extraction approach is used to analyze finger movement.

4.1 Experimental Data Set

In our finger movement classification analyzes, we obtained MI EEG signals recorded during imagination of the movements of the five fingers from an open available large electroencephalographic MI dataset. The 4 different MI task paradigms available in this dataset, we used 5F and NoMT paradigms 1 sec 21-channel EEG signals of 8 subjects. A total of six class categories are available in our classifications. 19 different EEG channels at sampling frequency of 1000 Hz are analyzed.

In the preprocessing section of finger movement classifications, 100 samples of 1000 Hz EEG signals of six different classes MIs (5F and NoMT paradigms) were selected to be analyzed in signal processing and following classification section. Hence, 600 EEG samples are used in signal processing.

The finger movement classification studies performed using six different feature sets and the ANOVA-based and PCA-based feature selection methods in this thesis are represented in Figure 4.1.

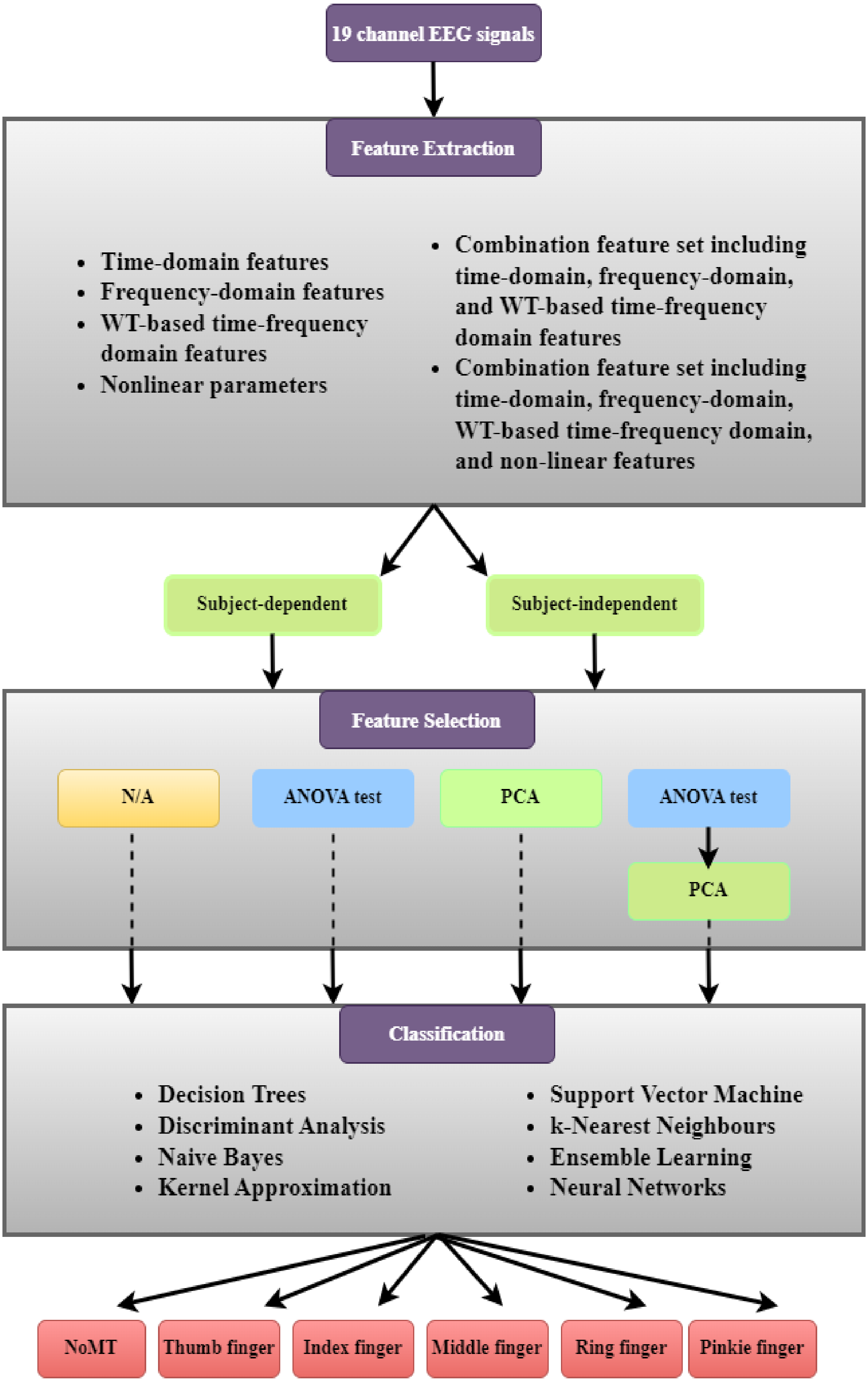


Figure 4.1: The block diagram of the proposed finger movement classification approach using different feature extraction approaches and feature selection methods with various classifiers.

4.2 Feature Extraction

24 time-domain, 15 frequency-domain, 15 WT-based time-frequency domain, 30 ITD-based time-frequency domain, and 4 non-linear features are evaluated for each EEG channel of each EEG segment. These feature sets and their different combination feature sets are analyzed using various machine learning algorithms for finger movement classifications, separately.

4.2.1 Time-domain Feature Set

After the extraction of EEG segments, the time-domain feature set was obtained evaluating 24 different amplitude and statistical information-based features in the time-domain. All time-domain features and the mathematical formulations of these features are given in Section 2.2.1. In the time-domain feature extraction-based approach, a total of 7800x456 and 600x456 size time-domain feature sets are created for subject-independent and subject-dependent finger movement classifications, respectively.

4.2.2 Frequency-domain Feature Set

To create the frequency-domain feature set, the frequency domain of EEG segments is generated using FT and five different EEG sub-bands are decomposed for each EEG segment. Energy, variance, and entropy measures are calculated using EEG sub-bands. The mathematical formulations of energy, variance and entropy values are available in Section 2.2.2. In the frequency-domain feature extraction-based approach, a total of 7800x285 and 600x285 size frequency-domain feature sets are generated for subject-independent and subject-dependent finger movement classifications, respectively.

4.2.3 Wavelet Transform-based Time-Frequency Domain Feature Set

To generate the WT-based time-frequency domain feature set, the time-frequency domain of EEG segments is obtained using WT and five different EEG sub-bands are decomposed for each EEG segment. Energy, variance, and entropy measures are calculated utilizing EEG sub-bands. Haar mother wavelet and 9 level sub-band decomposition are utilized for our finger movement classification analysis. The mathematical formulations of energy, entropy, and variance values based on the WT methos are given in Section 2.2.3.1. In the WT-based time-frequency domain feature extraction approach, a total of 7800x285 and 600x285 size time-frequency domain feature sets are created for subject-independent and subject-dependent finger movement classifications, respectively.

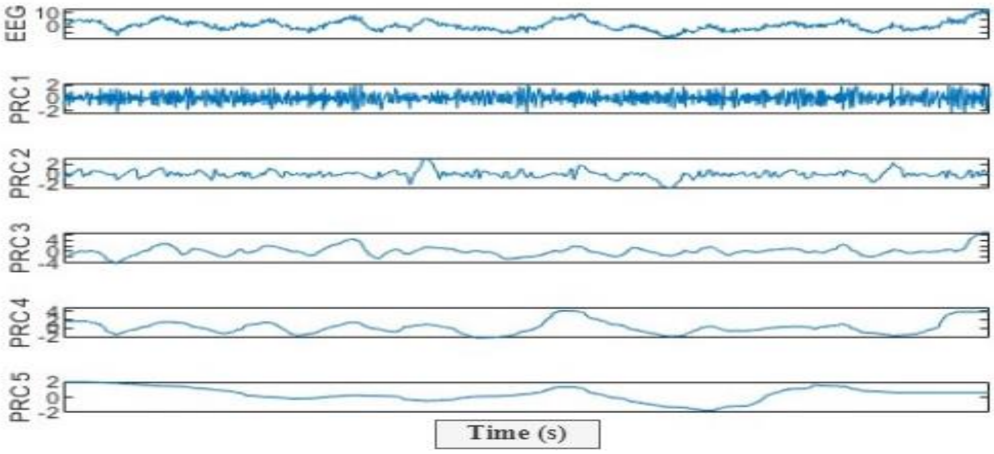
4.2.4 Intrinsic Time-Scale Decomposition-based Time-Frequency Domain Feature Set

In another time-frequency domain feature set, the features are evaluated using ITD algorithm. The different number of PRCs are obtained after applying the ITD algorithm. However, the defining of relevant PRC which best represents the EEG signal is an important task before feature extraction step. We performed energy-based feature selection process to define the best representative PRCs for feature extraction step. Firstly, the energies of each PRCs are evaluated as given in Equation (4.1).

$$Energy_{PRC_i} = \sum_{n=0}^{N-1} |PRC_i[n]|^2 , \qquad i = 1,\ 2,\ \dots ,\ L. \tag{4.1}$$

where, $Energy_{PRC_i}$is energy of ith PRC which is indicated as PRC_i. The first 5 PRCs of a 1 sec EEG signal and energies of them are given in Figure 4.2a and 4.2.b.

We selected the higher energy PRCs considering them as the best representative of the EEG signal. We observed that the energy of PRCs is decreased from PRC1 to PRC5. Therefore, the first 3 PRCs are used to extract features for our analysis due to their higher energy contents. We also investigated the effectiveness of different features. These are binary combinations (PRC1-PRC2, PRC1-PRC3 or PRC2-PRC3) and triple combination (PRC1-to-3) of these three PRCs. Then 10 time-frequency features which are power, mean value, sample entropy higher-order frequency moments (1st, 2nd, 3rd, and 4th moment), and Hjorth parameters (activity, mobility, and complexity) are evaluated utilizing the selected PRCs. For defining of the effect of ITD-based approach, the same features are evaluated from EEG segment itself, without the ITD application. The mathematical formulations of these time-frequency features based on the WT methos are given in Section 2.2.3.2.



(a) The extracted first five PRCs.

(b) The energies of these first five PRCs.

Figure 4.2: (a) The first 5 PRCs provided utilizing ITD, and (b) the energies of them.

In the ITD-based time-frequency domain feature extraction-based approach for subject-independent classifications, a total of 4800x190, 4800x380, and 4800x570 size time-frequency domain feature sets are obtained for the selected PRC (PRC1, PRC2 or PRC3), binary combinations of PRCs (PRC1-PRC2, PRC1-PRC3 or PRC2-PRC3), and triple combination (PRC1-to-3), respectively. In EEG-based analysis, a total of 4800x190 EEG feature set is obtained. In the ITD-based time-frequency domain feature extraction-based approach for subject-dependent classifications, a total of 600x190, 600x380, and 600x570 size time-frequency domain feature sets are obtained for the selected PRC (PRC1, PRC2 or PRC3), binary combinations of PRCs (PRC1-PRC2, PRC1-PRC3 or PRC2-PRC3), and triple combination (PRC1-to-3), respectively.

4.2.5 Non-linear Feature Set

Non-linear feature sets are provided using Poincare plot-based measures. 4 different non-linear features are evaluated to supply information about the non-linear dynamics embedded in EEG signals for each EEG segment where lag=1. Poincare plot measures' mathematical formulations are given in Section 2.2.4. In the non-linear feature extraction-based approach, a total of 7800x76 and 600x76 size non-linear feature sets are created for subject-independent and subject-dependent finger movement classification.

4.2.6 Combination Feature Set Including Time-domain, Frequency-domain, and Wavelet Transform-based Time-frequency domain Features

In addition to five different feature sets, the effectiveness of the combination of different feature sets are analyzed to improve the classification performance in our thesis studies. Combination feature sets are created using 24 time-domain, 15 frequency-domain, and 15 WT-based time-frequency domain features for each EEG channel of each EEG segment. In the combination feature set-based analysis, a total of 7800x1026 and 600x1026 size the combination feature sets are created for subject-independent and subject-dependent finger movement classifications, respectively.

4.2.7 Combination Feature Set Including Time-domain, Frequency-domain, Wavelet Transform-based Time-frequency domain, and Non-linear Features

In our second combination feature sets, we added 4 non-linear features into our first combination feature set which is created using 24 time-domain, 15 frequency-domain, and 15 WT-based time-frequency domain features for each EEG channel of each EEG segment. 4 non-linear features are evaluated for each EEG segment where lag=1. A total of 76 non-linear features are evaluated from all EEG channels of each EEG segment based on the lag=1 condition and added to our previous combination sets. In our second combination feature set-based analysis, a total of 7800x1102 and 600x1102 size the combination feature sets are created for subject-independent and subject-dependent finger movement classifications, respectively.

4.3 Feature Selection

In our finger movement classifications, we applied statistically significance-based feature selection method to improve the classifier performances selecting relevant and discriminative features. To compare the effectiveness of ANOVA, PCA-based feature selection method, which is generally utilized for the feature selection, was also used. Four different approaches are presented for the classification of each feature set.

Table 4.1: The number of features in all paradigms for time-domain feature set classifications.

Feature set	**All Features**	**PCA Selected Features**	**ANOVA Selected Features**	**ANOVA and PCA Selected Features**
S1	456	3	251	2
S2	456	1	262	1
S3	456	1	377	1
S4	456	1	383	1
S5	456	3	233	3
S6	456	3	264	2
S7	456	3	286	2
S8	456	5	192	1
All	456	5	318	4

According to our proposed feature selection methods, four different sets are created from our extracted feature sets to apply as input for classifiers. These are:

- All features in the corresponding feature set,
- PCA-selected principal components from the corresponding feature set,

- ANOVA-selected statistically significant features from the corresponding feature set,
- Both ANOVA and PCA selected features from the corresponding feature set.

Table 4.2: The number of features in all paradigms for frequency-domain feature set classifications.

Feature Set	All Features	PCA Selected Features	ANOVA Selected Features	ANOVA and PCA Selected Features
S1	285	3	117	2
S2	285	1	98	1
S3	285	1	154	1
S4	285	1	157	1
S5	285	3	119	3
S6	285	2	107	1
S7	285	3	116	1
S8	285	4	67	2
All	285	5	153	4

Table 4.3: The number of features in all paradigms for WT-based time-frequency domain feature set classifications.

Feature Set	All Features	PCA Selected Features	ANOVA Selected Features	ANOVA and PCA Selected Features
S1	285	1	10	1
S2	285	3	88	3
S3	285	2	39	2
S4	285	4	136	3
S5	285	1	25	1
S6	285	1	26	15
S7	285	1	135	1
S8	285	3	20	1
All	285	2	28	18

Table 4.4: The number of features in all paradigms for non-linear feature set classifications.

Feature Set	All Features	PCA Selected Features	ANOVA Selected Features	ANOVA and PCA Selected Features
S1	76	2	42	3
S2	76	3	45	3
S3	76	4	53	3
S4	76	3	63	3
S5	76	2	31	1
S6	76	4	33	3
S7	76	2	60	2
S8	76	3	32	2
All	76	4	38	2

Table 4.5: The number of features in all paradigms for combined (TD+FD+WT) feature set classifications.

Feature Set	All Features	PCA Selected Features	ANOVA Selected Features	ANOVA and PCA Selected Features
S1	1026	3	378	2
S2	1026	1	448	1
S3	1026	1	570	1
S4	1026	1	676	1
S5	1026	3	377	3
S6	1026	3	397	2
S7	1026	3	537	2
S8	1026	4	279	1
All	1026	5	499	4

The effectiveness of these four different feature sets is investigated and compared in all extracted feature sets using various classifiers. Table 4.1-4.6 summarizes the number of all features, ANOVA-selected, PCA-selected, and both ANOVA and PCA selected features in six different feature sets (TD, FD, WT, P, TD+FD+WT, and TD+FD+WT+P) which are used in our finger movement classifications.

Table 4.6: The number of features in all paradigms for combined (TD+FD+WT+P) feature set classifications.

Feature Set	All Features	PCA Selected Features	ANOVA Selected Features	ANOVA and PCA Selected Features
S1	1102	3	420	2
S2	1102	1	493	1
S3	1102	1	623	1
S4	1102	1	739	1
S5	1102	3	408	3
S6	1102	1	430	2
S7	1102	3	597	2
S8	1102	5	311	1
All	1102	5	537	3

Table 4.7: The number of features in all paradigms for ITD-based and EEG-based feature sets classifications.

Feature Set	PRC1, PRC2, PRC3, EEG	ANOVA + EEG	PRC1-PRC2, PRC1-PRC3, PRC2-PRC3	PRC1 to PRC3	ANOVA +PRC1 to PRC3
S1	190	108	380	570	180
S2	190	101	380	570	169
S3	190	147	380	570	319
S4	190	161	380	570	284
S5	190	127	380	570	193
S6	190	143	380	570	194
S7	190	101	380	570	255
S8	190	59	380	570	131
All	190	116	380	570	205

In our ITD-based finger movement classification, only ANOVA-based feature selection is used for triple combination (PRC1-to-PRC3) and EEG feature sets. Table 4.7 represents the number of features in different ITD-based feature sets and EEG-based feature sets for finger movement classification. In this table, PRC1, PRC2, PRC3 or EEG; indicate the number of features in feature set are obtained by using the corresponding PRC or EEG. ANOVA+EEG shows the ANOVA selected EEG features. PRC1-PRC2, PRC1-PRC3, and PRC2-PRC3 show the number of features in binary combination feature sets are extracted from PRC1 and PRC2, PRC1 and PRC3, and PRC2 and PRC3, respectively. "PRC1 to PRC3" indicates the number of features in triple combination feature set are calculated using all three PRCs. Additionally, "ANOVA+ PRC1 to PRC3" indicates the number of ANOVA-selected features in triple combination feature set are calculated using all three PRCs.

4.4 Results and Discussions of Finger Movement Classification

In this section, we show the performance results of finger movement classification provided by different feature extraction-based methods utilizing different machine learning algorithms. Seven different feature sets are created by various feature extraction methods using 1 sec finger movements EEG signals provided from an open-available EEG dataset. We calculated time-domain (TD), frequency-domain (FD), WT-based time-frequency domain (WT), ITD-based time-frequency domain, non-linear, and their two different combinations features sets using EEG segments which belongs to the six different classes (NoMT condition and 5 finger movements). We obtained 2 different combination feature sets. The first combination feature set (TD+FD+WT) includes time-domain, frequency-domain, and WT-based time-frequency features while the second combination feature set includes (TD+FD+WT+P) features of the first combination feature set and non-linear features. Additionally, we used two different feature selection methods such as ANOVA test and PCA to improve classifier performance defining relative features and reducing classifier complexity. These methods are applied to all feature sets except ITD-based time-frequency based feature set. We defined four different feature sets using ANOVA and PCA feature selection methods from our 6 different feature sets. These are categorized as (i) all feature set, (ii) PCA-selected feature set, (iii) ANOVA-selected feature set, and ANOVA and PCA-selected feature set from the corresponding feature set (TD, FD, WT, P, TD+FD+WT, and TD+FD+WT+P). These feature sets are classified using 8 different machine learning algorithms and accuracy-based performance evaluations are performed to investigate the effects of different feature sets and feature selection methods. In ITD-based approaches, 10 different features are calculated from the selected PRCs (PRC1, PRC2, and PRC3) provided by ITD and the EEG signal itself for each EEG segment. Three different PRCs (PRC1, PRC2, and PRC3), binary combinations of them (PRC1-PRC2, PRC1-PRC3, and PRC2-PRC3), and triple combination (PRC1-to-3). In addition, ANOVA test-based feature selection is performed and the effect on PRC1-to-3 and EEG feature sets is investigated. All feature sets are classified using DT, LDA, SVM, NB, k-NN, EL, NN, and KA algorithms and the results of all proposed approaches are analyzed based on the accuracy performance metric.

Table 4.8: Performance results (%) for finger movement classification using all features of time-domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	24.20	37.50	38.30	40.00	28.30	35.80	29.20	32.50	28.60
Discriminant Analysis	15.00	26.70	34.20	32.50	20.00	29.20	25.00	26.70	32.10
Naive Bayes	30.00	40.00	33.30	41.00	26.70	33.30	30.80	38.30	27.90
Support Vector Machine	28.30	**50.00**	**56.00**	48.30	**40.00**	45.80	28.30	**40.80**	**36.20**
k-Nearest Neighbours	**35.80**	44.20	45.00	42.00	34.20	45.00	27.50	38.30	33.50
Ensemble Learning	30.00	44.20	48.30	**50.00**	38.40	**46.70**	30.00	37.50	32.60
Neural Networks	29.20	45.00	51.50	47.50	35.80	45.80	**31.70**	38.30	34.90
Kernel Approximation	28.33	25.83	34.17	32.50	28.33	24.17	24.17	21.67	25.20

The performance results of our presented approaches using different feature sets such as time-domain, frequency-domain, WT-based time-frequency domain, non-linear feature sets and their two different combination feature sets (TD+FD+WT and

TD+FD+WT+P), two different feature selection methods (ANOVA test and PCA), and 8 different classifiers are given in Tables 4.8-4.31. In these tables, S1, S2, S3, S4, S5, S6, S7, or S8; indicate that features for subject-dependent classification are calculated utilizing the corresponding subject. "All subjects" indicates the features for subject-independent classification are calculated utilizing all subjects.

Table 4.9: Performance results (%) for finger movement classification using PCA-selected features of time-domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	27.50	**20.00**	27.50	32.50	22.50	20.00	26.70	19.20	23.00
Discriminant Analysis	25.80	18.30	27.00	32.50	20.80	**26.70**	18.30	18.30	20.90
Naive Bayes	23.30	18.30	28.30	**36.70**	21.70	22.50	25.80	16.70	21.90
Support Vector Machine	26.70	17.50	26.00	34.20	25.00	24.20	24.20	**23.30**	19.90
k-Nearest Neighbours	27.50	17.50	**31.00**	35.00	**27.50**	22.50	25.00	20.00	**24.00**
Ensemble Learning	**30.00**	**20.00**	27.00	34.20	25.00	**26.70**	**30.00**	19.20	23.00
Neural Networks	24.20	19.20	**31.00**	**36.70**	25.00	25.00	23.30	18.30	23.50
Kernel Approximation	27.50	16.70	17.00	16.70	25.80	17.50	21.70	22.50	21.40

Table 4.10: Performance results (%) for finger movement classification using ANOVA-selected features of time-domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	35.00	38.30	36.00	45.00	24.20	39.20	30.80	31.70	28.50
Discriminant Analysis	34.20	41.70	43.30	46.70	36.70	37.50	20.80	31.70	29.90
Naive Bayes	25.00	42.50	45.00	48.30	30.00	39.20	30.80	37.50	28.50
Support Vector Machine	35.00	**49.20**	**57.50**	54.20	39.20	**55.00**	33.30	**41.70**	**35.90**
k-Nearest Neighbours	29.20	45.00	47.00	50.00	35.00	48.30	30.80	39.20	32.70
Ensemble Learning	33.30	43.30	53.30	**55.80**	**43.30**	50.00	**34.20**	40.80	33.60
Neural Networks	**35.80**	46.70	57.00	50.80	40.80	48.30	**34.20**	40.00	34.70
Kernel Approximation	28.30	27.50	31.00	38.30	23.30	30.00	23.30	30.80	24.70

The performance results of all time-domain feature set-based classification with various classifiers are reported in Tables 4.8-4.11. The performance results show that SVM algorithm obtained 56.00% accuracy utilizing all time-domain features obtained from Subject C (S3). At the same time, the higher accuracy value (57.50%) of all time-domain based classifications is achieved using ANOVA-selected time-domain features of same subject and same classifier. In subject-independent analysis, the best result is achieved using all time-domain features and SVM classifier with accuracy of 36.20%. However, the accuracy of 35.90% is achieved using ANOVA-selected time-domain features and SVM algorithm. The results of all classifications performed using time-domain based approaches are summarized in Table 4.8-4.11.

Table 4.11: Performance results (%) for finger movement classification using both ANOVA and PCA selected features of time-domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	25.80	15.80	27.00	**43.30**	25.80	25.00	25.00	21.70	22.60
Discriminant Analysis	27.50	18.30	28.30	31.70	25.80	28.30	19.20	20.80	19.70
Naive Bayes	27.50	18.30	27.00	34.20	22.50	25.00	17.50	20.80	21.80
Support Vector Machine	25.80	**19.20**	**33.30**	33.30	27.50	30.00	22.50	20.00	20.00
k-Nearest Neighbours	**34.20**	16.70	29.20	40.80	28.30	**32.50**	**25.80**	**25.80**	22.80
Ensemble Learning	28.30	18.30	31.00	40.00	25.00	28.30	21.70	**25.80**	23.50
Neural Networks	30.00	17.50	31.00	35.80	**31.70**	27.50	24.20	20.80	**24.00**
Kernel Approximation	21.70	16.70	17.00	16.70	23.30	15.00	23.30	16.70	22.80

The selected statistically significant time-domain features distribution over 19 EEG channels was examined for subject-dependent and subject-independent finger movement classifications in Table 4.12 and Table 4.13. Firstly, in subject-independet finger movement classification, all time domain features except waveform length, average amplitude change value, absolute difference in standard deviation and slope-change value were mostly determined and selected as significant features by ANOVA in all channels. When examining the effectiveness of the channels, it was observed that statistically significant features were selected from all channels and did not concentrate on certain channels. However, selecting statistically significant features from all channels and specific time domain feature types did not provide improvement in classifier performance. It was observed that the highest performances in subject-independent analyzes were obtained with all time-domain features.

Table 4.12: Channel-based ANOVA-selected statistically significant feature distribution for subject-independent finger movement classification in time-domain feature set.

F_{id}	Channels																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	T
T_1																				18
T_2																				14
T_3																				18
T_4																				17
T_5																				15
T_6																				15
T_7																				11
T_8																				15
T_9																				15
T_{10}																				0
T_{11}																				0
T_{12}																				1
T_{13}																				11
T_{14}																				19
T_{15}																				15
T_{16}																				16
T_{17}																				16
T_{18}																				17
T_{19}																				19
T_{20}																				17
T_{21}																				14
T_{22}																				17
T_{23}																				17
T_{24}																				1
T	20	20	19	20	8	8	20	19	13	14	20	20	18	14	15	16	20	12	17	313

In subject-dependet finger movement classification, statistically significant features were selected intensively and balancedly from all channels and all feature types as can be seen from Table 4.13. In fact, it has been observed that feature selection that does not depend on a specific channel or feature group, performed with ANOVA, improves classifier performance. The highest classification performances in subject-dependent analyzes were obtained with ANOVA-selected time-domain features.

Table 4.13: Channel-based ANOVA-selected statistically significant feature distribution for subject-dependent finger movement classification in time-domain feature set.

F_{id}	Channels																			
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**T**
T_1	7	7	6	7	5	4	5	6	4	6	7	7	6	5	5	5	4	2	5	103
T_2	7	8	7	7	3	5	4	4	4	5	7	6	5	4	3	5	8	2	4	98
T_3	7	6	5	6	7	8	5	6	6	2	7	6	6	6	4	5	8	3	5	107
T_4	8	8	7	8	3	2	4	6	5	7	7	7	3	4	4	6	8	1	4	103
T_5	8	8	6	7	3	3	4	6	5	3	6	6	4	2	4	7	8	2	5	97
T_6	8	8	6	7	3	3	4	6	5	3	6	6	4	2	4	7	8	2	5	97
T_7	8	8	6	7	2	2	2	5	3	4	6	6	4	3	3	3	8	2	3	87
T_8	8	8	7	7	2	2	2	5	3	4	7	6	4	3	3	3	8	2	3	89
T_9	8	8	6	7	3	3	4	6	5	5	6	6	4	3	4	6	8	2	5	106
T_{10}	2	2	5	4	4	4	3	3	2	3	4	4	6	4	3	4	2	3	3	65
T_{11}	2	2	5	4	4	4	3	3	2	3	4	4	6	4	3	4	2	3	3	65
T_{12}	2	3	5	4	4	4	3	3	2	3	4	4	6	3	3	4	2	3	3	65
T_{13}	6	6	5	6	2	2	2	2	1	2	6	3	5	2	1	2	6	2	6	67
T_{14}	7	7	6	7	2	3	5	5	5	5	4	7	4	3	5	6	7	1	5	94
T_{15}	8	8	7	7	2	2	2	5	3	4	7	6	4	3	3	3	8	2	3	87
T_{16}	7	7	7	7	6	6	7	6	6	7	8	8	5	6	4	5	7	5	8	122
T_{17}	8	8	6	6	5	6	6	6	6	7	7	7	6	7	5	6	7	4	8	122
T_{18}	8	8	7	7	2	2	3	5	2	5	7	7	4	4	4	4	8	1	4	92
T_{19}	6	6	6	7	8	8	4	5	4	3	7	8	6	6	5	3	6	3	4	105
T_{20}	7	5	5	5	8	8	5	3	6	2	6	6	7	6	6	4	5	3	4	101
T_{21}	6	7	5	5	8	8	5	2	5	4	7	5	6	6	5	4	7	3	4	101
T_{22}	6	6	3	7	4	5	5	2	3	3	5	6	3	6	3	4	5	3	3	82
T_{23}	7	6	6	5	4	4	5	6	6	6	6	5	6	5	4	5	5	3	7	101
T_{24}	6	6	4	6	5	6	5	5	4	5	4	6	6	7	4	7	5	7	6	104
T	1-57	1-56	1-45	1-50	99	1-04	97	1-11	97	1-03	1-45	1-42	1-20	1-04	92	1-12	1-50	64	1-10	2258

The performance results of all frequency-domain feature set-based classification with various classifiers are reported in Tables 4.14-4.17. The performance results show that EL algorithm obtained 49.17% accuracy utilizing all frequency-domain features obtained from Subject C (S3). However, the higher accuracy value (55.00%) of all frequency-domain based classifications is achieved using ANOVA-selected frequency-domain features Subject E (S4) and same classifier. In subject-independent analysis, the best result is achieved using ANOVA-selected frequency-domain features and SVM classifier with accuracy of 30.45%.

Table 4.14: Performance results (%) for finger movement classification using all features of frequency-domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	29.17	30.83	35.83	35.83	34.17	34.17	34.67	23.33	24.10
Discriminant Analysis	25.00	30.00	40.00	39.17	28.33	35.83	24.17	23.33	28.14
Naive Bayes	25.00	29.17	35.83	37.50	25.00	36.67	25.00	27.50	23.91
Support Vector Machine	28.33	**39.17**	40.83	40.00	34.17	35.00	29.17	30.83	**29.42**
k-Nearest Neighbours	26.67	30.83	38.33	36.67	30.83	32.50	29.17	28.33	24.62
Ensemble Learning	**30.83**	38.33	**49.17**	41.67	**40.00**	**41.67**	**36.67**	28.33	28.21
Neural Networks	28.33	34.17	43.33	**44.17**	33.33	37.50	29.17	30.00	27.69
Kernel Approximation	25.00	20.00	37.50	40.00	24.17	25.83	22.50	**33.33**	25.71

Table 4.15: Performance results (%) for finger movement classification using PCA-selected features of frequency-domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	30.00	16.17	**31.67**	30.83	28.33	21.67	23.33	27.50	23.91
Discriminant Analysis	26.67	**20.00**	20.00	30.83	20.83	24.17	19.17	27.50	23.14
Naive Bayes	28.33	19.17	30.83	30.00	**30.83**	**25.00**	23.33	22.50	22.69
Support Vector Machine	27.50	17.50	28.33	31.67	25.83	**25.00**	22.50	26.67	21.86
k-Nearest Neighbours	28.33	19.17	29.17	32.50	28.33	24.17	21.67	**30.00**	24.49
Ensemble Learning	26.67	20.00	30.00	30.83	23.33	23.33	20.83	27.50	24.68
Neural Networks	**34.17**	17.50	29.17	**36.67**	29.17	22.50	**24.17**	29.17	**25.83**
Kernel Approximation	**34.17**	16.67	16.67	16.67	24.17	16.67	22.50	25.00	21.86

Table 4.16: Performance results (%) for finger movement classification using ANOVA-selected features of frequency-domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	29.17	32.50	33.33	33.33	36.67	25.83	**34.17**	30.00	25.90
Discriminant Analysis	25.83	35.00	48.33	51.67	41.67	32.50	26.67	31.67	26.60
Naive Bayes	28.33	30.83	40.83	47.50	31.67	27.50	25.83	24.17	25.00
Support Vector Machine	30.00	45.00	50.00	54.17	40.83	**40.83**	29.17	31.67	**30.45**
k-Nearest Neighbours	**34.17**	34.17	37.50	45.00	35.83	32.50	29.17	29.17	26.09
Ensemble Learning	31.67	**45.83**	**51.67**	**55.00**	**47.50**	37.50	33.33	29.17	28.85
Neural Networks	29.17	40.00	50.00	51.67	38.33	38.33	30.00	**35.83**	27.63
Kernel Approximation	26.67	28.33	37.50	40.83	27.50	23.33	25.83	26.67	27.05

Table 4.17: Performance results (%) for finger movement classification using both ANOVA and PCA selected features of frequency-domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	27.50	20.00	28.33	**37.50**	26.67	25.00	20.83	25.00	24.74
Discriminant Analysis	28.33	20.00	26.67	34.17	23.33	**29.17**	19.17	20.83	21.54
Naive Bayes	29.17	19.17	**32.50**	**37.50**	**30.00**	22.50	18.33	20.83	24.04
Support Vector Machine	30.00	18.33	31.67	33.33	26.67	25.83	22.50	24.17	21.60
k-Nearest Neighbours	30.00	17.50	31.67	35.00	27.50	22.50	22.50	25.00	24.10
Ensemble Learning	28.33	**22.50**	25.83	35.00	28.33	**29.17**	20.83	23.33	23.33
Neural Networks	**33.33**	17.50	30.00	**37.50**	**30.00**	24.17	**25.00**	**30.00**	**25.12**
Kernel Approximation	22.50	16.67	16.67	16.67	24.17	16.67	16.67	18.33	23.08

The selected statistically significant frequency-domain features distribution over 19 EEG channels was investigated for subject-dependent and subject-independent finger movement classifications in Table 3.18 and Table 3.19. For subject-independet finger movement classification, it has been observed that in selecting statistically significant features, different features are focused on in different EEG frequency bands and the same features are not indicated as statistically significant features in each frequency band. When the statistically significant feature distribution in the channels was examined, it was seen that balanced selections were made from all channels. The highest accuracy value of subject-independent classification was achieved using ANOVA-selected frequency domain features. As a result, using ANOVA, classifier performance was improved by selecting statistically significant features from all channels and features, rather than selecting features by focusing on specific channels and features.

Table 4.18: Channel-based ANOVA-selected statistically significant feature distribution for subject-independent finger movement classification in frequeny-domain feature set.

F_{id}	Channels																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**T**
F_1	■	■	■	■			■	■			■	■	■		■	■	■	■	■	14
F_2	■	■	■	■	■	■	■	■			■	■	■	■	■	■	■	■	■	17
F_3	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	■	■	18
F_4	■	■	■	■	■	■	■	■			■	■	■	■		■	■	■	■	16
F_5	■	■	■	■							■	■	■				■	■	■	10
F_6				■									■							2
F_7		■	■		■	■	■	■	■	■	■		■	■		■	■	■		14
F_8	■	■	■		■	■	■	■	■	■			■	■		■		■		13
F_9					■	■										■				3
F_{10}	■	■	■	■			■	■	■	■	■	■	■	■		■	■	■	■	16
F_{11}		■	■				■	■	■	■	■	■		■			■	■	■	12
F_{12}									■	■					■					3
F_{13}			■								■			■						3
F_{14}			■								■			■						3
F_{15}					■	■		■		■	■		■	■			■	■		9
T	7	9	11	7	7	7	8	9	6	7	11	7	10	9	4	8	9	10	7	153

In subject-dependet finger movement classification, statistically significant frequency-domain features were indicated and selected intensively and balancedly from all channels and 15 different features. In fact, it has been observed that feature selection that does not depend on a specific channel or feature group, performed with ANOVA, improves classifier performance. The highest classification performances in subject-dependent analyzes were obtained with ANOVA-selected frequency-domain features.

Table 4.19: Channel-based ANOVA-selected statistically significant feature distribution for subject-dependent finger movement classification in frequeny-domain feature set.

F_{id}	Channels																			
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**T**
F_1	8	8	8	7	3	3	4	5	4	3	5	8	4	2	5	4	8	3	6	98
F_2	8	8	8	7	3	3	3	5	4	2	4	7	3	3	5	4	8	4	5	94
F_3	6	4	5	4	2	2	4	3	3	2	4	4	2	3	5	3	3	4	4	68
F_4	6	7	3	5	3	2	4	5	3	3	2	6	3	5	5	4	4	7	6	81
F_5	6	6	2	5			2	3		1	1	5		4	2	1	3	5		46
F_6	2	2		2					1	1		1	1	1			1		1	13
F_7	3	3	1	2	6	5	4	4	2	6	3	4	4	4	4	5	2	4	5	71
F_8	5	5	2	1	5	5	4	4	2	3	2	5	3	4	5	4	2	2	3	66
F_9	1	2			1	3	2	1				1	1	2		2				16
F_{10}	5	4	6	5	4	4	3	3	3	4	6	6	7	7	2	4	4	3	3	83
F_{11}	6	5	3	4	3	3	3	3	5	5	5	3	5	6	2	3	3	3	3	73
F_{12}	4	4		1			2	3	1	2	2	3	1	2	1	3	2	3	2	36
F_{13}	5	5	5	1	2	2		4	2	4	4	5	6	5	2	3	3	1	1	60
F_{14}	5	5	4	2	2	3		4	2	3	4	5	6	5	1	3	6	2	2	64
F_{15}	3	4	3	4	4	4	4	4	3	3	3			4	2	3	6	4	5	63
T	73	72	50	50	38	39	39	54	35	42	46	63	47	57	41	46	55	45	53	9-32

Table 4.20: Performance results (%) for finger movement classification using all features of WT-based time-frequency domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	29.17	30.83	26.67	35.00	22.50	31.67	26.67	20.83	22.44
Discriminant Analysis	17.50	19.17	**31.67**	30.83	15.83	30.83	17.50	23.33	22.12
Naive Bayes	29.17	34.17	25.00	31.67	24.17	29.17	24.17	21.67	21.54
Support Vector Machine	32.50	**37.50**	30.83	29.17	25.83	30.83	26.67	30.00	22.00
k-Nearest Neighbours	26.67	34.17	27.50	27.50	24.17	**33.33**	29.17	**30.83**	22.12
Ensemble Learning	**35.00**	33.33	28.33	**35.83**	27.50	31.67	**33.33**	25.83	**26.60**
Neural Networks	28.33	30.83	30.83	30.00	18.33	31.67	25.83	25.83	21.22
Kernel Approximation	27.50	28.33	**31.67**	29.17	**32.50**	29.17	**33.33**	21.67	26.54

The performance results of all WT-based time-frequency domain feature set-based classification with various classifiers are reported in Tables 4.20-4.23. The performance results reveal that SVM algorithm provided 34.17% accuracy utilizing all time-frequency domain features obtained from Subject E (S4). However, the higher accuracy value (36.67%) of all WT-based time-frequency domain-based classifications is achieved using ANOVA-selected time-frequency domain features Subject E (S4) and both SVM and EL classifiers. In subject-independent analysis, the best result is achieved using all time-frequency domain features and EL algorithm with accuracy of 26.60%. On the other hand, 21.28% accuracy was achieved using the ANOVA-selected time-frequency features by the SVM algorithm.

Table 4.21: Performance results (%) for finger movement classification using PCA-selected features of WT-based time-frequency domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	20.00	19.17	25.00	31.67	29.17	23.33	30.83	25.83	20.71
Discriminant Analysis	16.67	20.83	20.83	28.33	15.00	17.50	16.67	18.33	17.00
Naive Bayes	18.33	22.50	26.67	26.67	23.33	**30.00**	30.83	19.17	19.68
Support Vector Machine	18.33	**25.00**	20.83	**34.17**	17.50	21.67	25.83	22.50	17.44
k-Nearest Neighbours	20.00	23.33	22.50	32.50	**31.67**	25.83	30.83	26.67	20.58
Ensemble Learning	**21.67**	21.67	**28.33**	30.00	30.00	23.33	30.00	25.83	19.36
Neural Networks	19.17	20.00	**28.33**	30.83	**31.67**	26.67	33.33	**27.50**	**21.09**
Kernel Approximation	16.67	12.50	22.50	25.00	16.67	16.67	**16.67**	21.67	17.05

Table 4.22: Performance results (%) for finger movement classification using ANOVA-selected features of WT-based time-frequency domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	20.83	24.17	27.50	33.33	24.17	22.50	29.17	25.00	19.68
Discriminant Analysis	24.17	17.50	30.83	32.50	25.00	30.00	24.17	22.50	20.77
Naive Bayes	**26.67**	30.00	31.67	33.33	24.17	31.67	25.00	25.00	19.81
Support Vector Machine	25.83	**31.67**	**33.33**	**36.67**	**25.83**	**32.50**	27.50	26.67	**21.28**
k-Nearest Neighbours	24.17	**31.67**	**33.33**	32.50	25.00	30.83	28.33	**29.17**	20.71
Ensemble Learning	25.83	30.83	30.83	**36.67**	**25.83**	31.67	29.17	24.17	20.77
Neural Networks	22.50	24.17	26.67	31.67	21.67	25.83	23.33	27.50	20.71
Kernel Approximation	18.33	23.33	26.67	23.33	16.67	31.67	**33.33**	22.50	19.55

Table 4.23: Performance results (%) for finger movement classification using both ANOVA and PCA selected features of WT-based time-frequency domain feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	26.67	**25.83**	28.33	31.67	22.50	28.33	25.00	**26.67**	19.42
Discriminant Analysis	25.83	21.67	29.17	27.50	24.17	**31.67**	18.33	24.17	20.45
Naive Bayes	28.33	**25.83**	25.00	29.17	20.00	25.83	25.83	22.50	**21.47**
Support Vector Machine	25.83	25.00	**32.50**	30.00	20.00	30.00	23.33	24.17	20.32
k-Nearest Neighbours	**29.17**	**25.83**	30.00	**32.50**	22.50	30.00	29.17	21.67	20.06
Ensemble Learning	25.83	20.83	29.17	29.17	24.17	**31.67**	26.67	24.17	20.58
Neural Networks	26.67	25.00	30.83	30.83	**25.83**	29.17	**30.83**	24.17	20.21
Kernel Approximation	16.67	19.17	20.83	25.00	16.67	28.33	16.67	16.67	19.36

Table 4.24: Channel-based ANOVA-selected statistically significant feature distribution for subject-independent finger movement classification in WT-based time-frequeny domain feature set.

F_{id}	Channels																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**T**
W_1																				0
W_2																				0
W_3																				9
W_4																				0
W_5																				0
W_6																				8
W_7																				0
W_8																				0
W_9																				6
W_{10}																				0
W_{11}																				0
W_{12}																				6
W_{13}																				0
W_{14}																				0
W_{15}																				1
T	0	4	3	0	0	0	3	0	1	0	3	1	4	4	5	0	0	0	0	28

The selected statistically significant WT-based time-frequency domain features distribution over 19 EEG channels was investigated for subject-dependent and subject-independent finger movement classifications in Table 3.24 and Table 3.25. For subject-independet finger movement classification, it has been observed that in selecting statistically significant features with ANOVA, entropy values are focused on in different frequency bands such as delta, theta, alpha, and beta EEG subbands. Among the 19 EEG channels, it was observed that no statistically significant features were selected from some channels. Although more statistically significant feature selections were made from certain features and channels in feature selection with ANOVA, it could not improve the classifier performance. The highest accuracy value of subject-independent classification was obtained using all WT-based time-frequency features.

In subject-dependet finger movement classification, statistically significant WT-based time-frequency features were indicated and selected intensively and balancedly from 19 EEG channels and 15 different features. In fact, it has been observed that feature selection that does not depend on a specific channel or feature group, performed with ANOVA, improves classifier performance. The highest classification performances in subject-dependent analyzes were obtained with ANOVA-selected WT-based time-frequency features.

Table 4.25: Channel-based ANOVA-selected statistically significant feature distribution for subject-dependent finger movement classification in WT-based time-frequeny domain feature set.

F_{id}	**Channels**																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**T**
W_1	3	1	1	1	2	2	1	2	3	3	1		4	1	1	2	3	1	3	35
W_2	3	1	1	1	2	2	1	2	3	3	1		4	1	1	2	3	1	3	35
W_3	1	5	2		2	1	3	1	1	1	2	1	6	3	2	1	3		2	37
W_4	3	1	1	1	2	2	1	2	3	3	1		4	1	1	2	3	1	3	35
W_5	3		1	1	2	2	1	2	3	2	1		4	1	1	2	3	1	3	33
W_6	3	4	1		3	2	3	2	3	3	2	2	6	4	3	1	4		1	47
W_7	3	2		1	2	2		2	3	3	1	1	4	2		2	3	1	2	34
W_8	3	2		1	2	2		2	3	3	1	1	3	2		2	3	1	2	33
W_9	1	5		2		2	1	4	2	3	1	2	3	4	2	2	4	2	2	42
W_{10}	3	1		1	2	1		2	2	2	1	1	4	2	1	3	3	1	2	32
W_{11}	3	1		1	2	1		2	2	2	1	1	4	2		3	3	1	2	31
W_{12}			2	1	1		1		1		1	2	2	2	2		3			18
W_{13}	3	1		2	1	1		2	1	1	2	1	3	2	1	2	3	2	2	30
W_{14}	3	1		2	1	1		2	1	1	1	1	2	2	1	2	3	2	2	28
W_{15}	2	1	1							1		1	1	1	2		1			11
T	37	26	10	15	24	21	12	27	31	31	17	14	54	30	18	26	45	14	29	4-81

Table 4.26: Performance results (%) for finger movement classification using all features of non-linear feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	29.17	30.83	29.17	34.17	28.33	34.17	**33.33**	24.17	25.26
Discriminant Analysis	25.00	33.33	41.67	39.17	30.83	35.00	26.67	34.17	27.24
Naive Bayes	30.83	32.50	30.83	33.33	25.83	31.67	30.00	27.50	23.14
Support Vector Machine	32.50	**39.17**	40.00	**43.33**	**37.50**	35.83	30.00	**35.83**	**30.90**
k-Nearest Neighbours	28.33	32.50	36.67	36.67	32.50	35.00	28.33	33.33	30.64
Ensemble Learning	30.83	36.67	44.17	38.33	**37.50**	35.83	30.00	33.33	29.81
Neural Networks	**31.67**	35.00	**45.00**	35.83	30.83	**36.67**	26.67	30.83	29.36
Kernel Approximation	26.67	28.33	32.50	31.67	24.17	28.33	28.33	24.17	26.09

Table 4.27: Performance results (%) for finger movement classification using PCA-selected features of non-linear feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	26.67	23.33	28.33	36.67	25.00	31.67	21.67	21.67	23.85
Discriminant Analysis	25.00	21.67	27.50	34.17	21.67	33.33	23.33	**33.33**	19.29
Naive Bayes	29.17	22.50	25.83	36.67	**28.33**	28.33	25.00	27.50	22.63
Support Vector Machine	26.67	23.33	**33.33**	**39.17**	20.83	33.33	24.17	25.83	21.92
k-Nearest Neighbours	28.33	**26.67**	29.17	35.83	26.67	**36.67**	**27.50**	26.67	24.81
Ensemble Learning	**30.83**	24.17	26.67	36.67	**28.33**	32.50	24.17	28.33	22.69
Neural Networks	28.33	**26.67**	28.33	35.00	25.00	32.50	26.67	26.67	**25.06**
Kernel Approximation	20.00	21.67	26.67	**39.17**	22.50	26.67	18.33	18.33	22.56

The performance results of all non-linear feature set-based classification with various classifiers are reported in Tables 4.26-4.29. The performance results reveal that NN algorithm achieved 45.00% accuracy utilizing all non-linear features obtained from Subject C (S3). However, the higher accuracy value (50.00%) of all non-linear feature set-based classifications is achieved using ANOVA-selected non-linear features Subject E (S4) and SVM classifier. In subject-independent analysis, the best result is achieved using ANOVA-selected non-linear features and SVM algorithm with accuracy of 31.79%. On the other hand, 30.90% accuracy is achieved using all non-linear features by the SVM algorithm.

Table 4.28: Performance results (%) for finger movement classification using ANOVA-selected features of non-linear feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	**34.17**	25.83	29.17	41.67	23.33	30.83	**33.33**	28.33	24.42
Discriminant Analysis	30.00	37.50	**45.00**	46.67	32.50	33.33	23.33	30.00	27.05
Naive Bayes	28.33	33.33	30.83	42.50	29.17	35.00	**33.33**	28.33	21.73
Support Vector Machine	**34.17**	**38.33**	43.33	**50.00**	35.00	34.17	29.17	33.33	**31.79**
k-Nearest Neighbours	31.67	35.83	33.33	43.33	30.83	32.50	30.83	28.33	28.27
Ensemble Learning	30.00	36.67	39.17	45.83	**40.83**	35.83	30.83	30.83	27.69
Neural Networks	30.83	37.50	40.00	42.50	31.67	**38.33**	30.83	**35.00**	29.62
Kernel Approximation	24.17	30.00	33.33	38.33	17.50	28.33	29.17	27.50	26.54

Table 4.29: Performance results (%) for finger movement classification using both ANOVA and PCA selected features of non-linear feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	28.33	20.83	35.00	34.17	25.83	21.67	24.17	20.83	24.04
Discriminant Analysis	30.00	21.67	30.83	34.17	20.83	27.50	19.17	**30.83**	20.58
Naive Bayes	27.50	20.83	35.83	36.67	26.67	**30.83**	22.50	29.17	21.73
Support Vector Machine	30.83	22.50	38.33	**39.17**	20.83	25.83	23.33	30.00	21.79
k-Nearest Neighbours	29.17	**27.50**	38.33	36.67	26.67	29.17	31.67	25.83	**24.68**
Ensemble Learning	**31.67**	24.17	35.00	38.33	23.33	28.33	30.83	**30.83**	24.04
Neural Networks	**31.67**	25.83	**40.00**	35.00	**27.50**	24.17	**32.50**	29.17	23.72
Kernel Approximation	24.17	25.00	35.83	38.33	16.67	25.83	23.33	13.33	21.35

The selected statistically significant nonlinear domain features distribution over 19 EEG channels was examined for subject-dependent and subject-independent finger movement classifications in Table 3.30 and Table 3.31. For subject-independet finger movement classification, it has been observed that in selecting statistically significant features with ANOVA, SD_2 and SD_1/SD_2 values where lag=1 were mostly selected as statistically significant features in most of the channels. Among 19 EEG channels, the distribution of statistically significant non-linear feature is balanced. With anova, balanced statistically significant feature distribution on these two non-linear features and in all channels increased the classifier performance. The highest accuracy value of subject-independent classification was obtained using ANOVA-selected non-linear features.

Table 4.30: Channel-based ANOVA-selected statistically significant feature distribution for subject-independent finger movement classification in non-linear domain feature set.

F_{id}	Channels																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	T
P_1																				1
P_2																				17
P_3																				4
P_4																				16
T	3	3	2	3	0	0	2	2	2	2	2	2	2	1	3	2	3	2	2	38

Table 4.31: Channel-based ANOVA-selected statistically significant feature distribution for subject-dependent finger movement classification in non-linear domain feature set.

F_{id}	Channels																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	T
P_1	3	3	5	3	3	3	2	4	3	4	4	3	6	4	3	4	3	2	4	66
P_2	8	8	7	8	1	1	3	5	4	6	6	7	4	4	4	5	8	2	5	96
P_3	6	6	5	5	1	1	3	3	2	4	6	4	4	3	2	2	6	3	4	70
P_4	7	7	6	6	6	6	8	7	6	7	8	8	6	7	5	6	7	6	8	127
T	24	24	23	22	11	11	16	19	15	21	24	22	20	18	14	17	24	13	21	359

In subject-dependet finger movement classification, statistically significant non-linear features were indicated and selected intensively and balancedly from 19 EEG channels and 4 different features. The feature selection that does not depend on a specific channel or feature group, performed with ANOVA, improves classifier performance. The highest classification performances in subject-dependent analyzes were obtained with ANOVA-selected non-linear features.

Table 4.32: Performance results (%) for finger movement classification using all features of combined (TD+FD+WT) feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	30.00	36.67	35.00	43.33	32.50	33.33	29.17	34.17	31.20
Discriminant Analysis	26.67	28.33	44.17	42.50	27.50	32.50	24.17	24.17	32.40
Naive Bayes	28.33	33.33	38.33	42.50	29.17	29.17	20.83	32.50	26.20
Support Vector Machine	33.33	**50.00**	**57.50**	51.67	**39.17**	45.00	28.33	**42.50**	**37.00**
k-Nearest Neighbours	**36.67**	38.33	49.17	40.83	34.17	39.17	31.67	34.17	32.30
Ensemble Learning	31.67	41.67	44.17	53.33	33.33	43.33	28.33	35.83	34.70
Neural Networks	32.50	44.17	53.33	**55.00**	**39.17**	**50.00**	**32.50**	39.17	34.70
Kernel Approximation	30.00	24.17	28.33	43.33	28.33	33.33	25.83	25.00	25.40

Table 4.33: Performance results (%) for finger movement classification using PCA-selected features of combined (TD+FD+WT) feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	25.00	19.17	30.83	35.83	26.67	**30.83**	20.83	24.17	22.80
Discriminant Analysis	**30.00**	**23.33**	27.50	35.00	25.00	29.17	19.17	21.67	20.90
Naive Bayes	23.33	**23.33**	26.67	35.00	20.83	27.50	25.00	23.33	22.10
Support Vector Machine	27.50	19.17	**31.67**	33.33	25.00	28.33	23.33	21.67	19.60
k-Nearest Neighbours	25.83	19.17	28.33	**37.50**	**28.33**	26.67	**27.50**	**26.67**	**24.00**
Ensemble Learning	30.00	21.67	27.50	36.67	27.50	29.17	20.00	24.17	22.80
Neural Networks	25.00	13.33	28.33	30.83	27.50	27.50	20.83	23.33	22.80
Kernel Approximation	22.50	16.67	16.67	16.67	20.83	25.83	24.17	21.67	22.10

Table 4.34: Performance results (%) for finger movement classification using ANOVA-selected features of combined (TD+FD+WT) feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	29.17	35.00	35.00	44.17	28.33	33.33	25.00	30.00	30.30
Discriminant Analysis	29.17	18.33	25.83	38.33	32.50	35.83	15.00	31.67	34.20
Naive Bayes	27.50	34.17	36.67	40.83	31.67	31.67	30.00	30.83	27.10
Support Vector Machine	35.83	**55.83**	55.00	50.00	39.17	**48.33**	**33.33**	37.50	**38.70**
k-Nearest Neighbours	29.17	45.00	45.00	41.67	35.83	41.67	**33.33**	31.67	32.90
Ensemble Learning	31.67	53.33	51.67	50.83	42.50	46.67	29.17	**45.83**	35.50
Neural Networks	**37.50**	54.17	**55.83**	**54.17**	**43.33**	47.50	30.83	35.00	36.10
Kernel Approximation	26.67	29.17	26.67	31.67	26.67	25.83	20.00	20.83	26.10

Table 4.35: Performance results (%) for finger movement classification using both ANOVA and PCA selected features of combined (TD+FD+WT) feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	27.50	15.83	26.67	34.17	20.83	29.17	25.00	21.67	23.40
Discriminant Analysis	28.33	18.33	**29.17**	**36.67**	20.00	35.00	24.17	20.83	19.80
Naive Bayes	28.33	18.33	27.50	**36.67**	22.50	33.33	20.00	20.83	21.40
Support Vector Machine	27.50	**22.50**	27.50	34.17	**25.83**	28.33	22.50	20.00	20.90
k-Nearest Neighbours	**31.67**	16.67	28.33	35.00	25.00	30.00	**26.67**	**25.83**	21.80
Ensemble Learning	28.33	18.33	**29.17**	35.83	20.83	35.83	25.00	**25.83**	23.40
Neural Networks	29.17	19.17	25.83	**36.67**	25.00	**37.50**	21.67	20.00	**24.50**
Kernel Approximation	18.33	16.67	16.67	16.67	23.33	15.00	20.00	16.67	22.50

In order to analyze the effect of different feature sets on finger movements classification, and compare these approaches, we investigated the combination feature set including time-domain, frequency-domain, and WT-based time-frequency features. The performances of classification performed using this combination set are given in Tables 4.32-4.35. The higher accuracy value (57.50%) of all combination set-based classifications is achieved all features of the combination set obtained from Subject C (S3) and SVM classifier. However, ANOVA-selected features of Subject B (S2) with SVM algorithm and Subject C (S3) with NN algorithm yielded accuracy of 55.83%. In subject-independent analysis, the best result is achieved using ANOVA-selected the combination set features and SVM algorithm with accuracy of 38.70%. On the other hand, 37.00% accuracy is achieved using all the combination set features by the SVM algorithm. The results of all classifications performed using the combination set-based approaches are provided in Table 4.32-4.35.

Table 4.36: Performance results (%) for finger movement classification using all features of combined (TD+FD+WT+P) feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	28.33	33.33	35.00	36.67	36.67	37.50	30.83	35.83	30.30
Discriminant Analysis	25.83	31.67	43.33	38.33	20.00	37.50	24.17	24.17	32.50
Naive Bayes	26.67	34.17	36.67	35.83	27.50	35.00	**31.67**	33.33	27.10
Support Vector Machine	30.00	**48.33**	55.00	50.00	**38.33**	42.50	26.67	**41.67**	**37.60**
k-Nearest Neighbours	29.17	43.33	47.50	45.00	30.83	39.17	29.17	32.50	32.10
Ensemble Learning	27.50	40.83	41.67	55.00	36.67	42.50	28.33	40.83	36.20
Neural Networks	**31.67**	42.50	**55.83**	**55.83**	35.83	**44.17**	27.50	40.83	34.40
Kernel Approximation	28.33	29.17	28.33	36.67	30.00	23.33	20.83	25.83	26.00

We investigated the effect of another combination, which is denoted as TD+FD+WT+P, including our previous combination set features with non-linear features. The performances of classification performed using this combination set are given in Tables 4.36-4.39. The higher accuracy value (59.17%) of all combination set-based classifications is achieved ANOVA-selected features of the combination set obtained from Subject E (S4) and SVM classifier. However, all features of combination set of Subject C (S3) and Subject E (S4) with NN algorithm yielded accuracy of 55.83%. In subject-independent analysis, the best result is achieved using

Table 4.37: Performance results (%) for finger movement classification using PCA-selected features of combined (TD+FD+WT+P) feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	28.33	17.50	30.83	37.50	26.67	**28.33**	26.67	16.67	22.90
Discriminant Analysis	28.33	20.83	27.50	34.17	25.00	24.17	18.33	20.00	19.60
Naive Bayes	30.83	**23.33**	26.67	34.17	20.83	25.00	25.83	21.67	21.50
Support Vector Machine	26.67	20.00	**31.67**	33.33	24.17	**28.33**	24.17	20.00	19.60
k-Nearest Neighbours	**31.67**	**23.33**	28.33	38.33	**28.33**	25.00	25.00	**25.00**	24.30
Ensemble Learning	30.83	20.83	27.50	35.00	27.50	24.17	**30.00**	21.67	23.30
Neural Networks	30.83	22.50	27.50	**41.67**	26.67	25.83	22.50	**25.00**	**24.70**
Kernel Approximation	28.33	16.67	16.67	16.67	21.67	16.67	19.17	21.67	22.60

Table 4.38: Performance results (%) for finger movement classification using ANOVA-selected features of combined (TD+FD+WT+P) feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	30.00	41.67	36.67	40.00	30.83	36.67	30.83	35.00	30.40
Discriminant Analysis	25.00	15.83	35.83	35.00	36.67	24.17	21.67	31.67	34.40
Naive Bayes	28.33	39.17	35.00	43.33	31.67	31.67	31.67	35.00	26.90
Support Vector Machine	**36.67**	**46.67**	**56.67**	**59.17**	41.67	**51.67**	32.50	37.50	**39.30**
k-Nearest Neighbours	31.67	41.67	43.33	45.00	29.17	45.83	31.67	34.17	33.30
Ensemble Learning	N/A	45.00	50.00	52.50	42.50	44.17	**34.17**	**40.83**	35.80
Neural Networks	N/A	**46.67**	55.83	57.50	**45.00**	49.17	30.00	**40.83**	37.20
Kernel Approximation	N/A	25.00	24.17	38.33	23.33	27.50	22.50	20.83	26.00

ANOVA-selected the combination set features and SVM algorithm with accuracy of 39.30%. On the other hand, 37.60% accuracy is achieved using all features of the combination set by the SVM algorithm. The results of all classifications performed using the combination set-based approaches are provided in Table 4.36-4.39. The results of two different feature sets reveal that the success of finger movement classification improved with including of non-linear features especially in ANOVA-selected features-based approaches.

Table 4.39: Performance results (%) for finger movement classification using both ANOVA and PCA selected features of combined (TD+FD+WT+P) feature set.

Models	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
Decision Tree	25.00	14.17	27.50	**43.33**	20.83	23.33	25.00	17.50	22.60
Discriminant Analysis	30.83	**18.33**	27.50	31.67	20.00	**28.33**	24.17	18.33	19.70
Naive Bayes	26.67	**18.33**	29.17	34.17	22.50	21.67	17.50	20.00	21.80
Support Vector Machine	28.33	**18.33**	**30.00**	35.00	**25.83**	25.83	22.50	21.67	19.90
k-Nearest Neighbours	25.83	17.50	**30.00**	40.83	25.00	24.17	**25.83**	**23.33**	22.80
Ensemble Learning	**31.67**	**18.33**	27.50	40.00	20.83	**28.33**	21.67	22.50	**23.50**
Neural Networks	29.17	17.50	29.17	38.33	24.17	25.83	24.17	20.00	23.40
Kernel Approximation	17.50	16.67	16.67	16.67	20.00	23.33	16.67	16.67	22.80

Table 4.40: Finger movement classification performance (%) of ITD based feature sets using the Decision Tree classifier.

Components	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
PRC1	29.17	27.50	35.00	32.50	**29.17**	25.83	31.67	27.50	**25.83**
PRC2	26.67	28.33	26.67	30.00	26.67	**30.83**	30.00	**32.50**	20.63
PRC3	24.17	28.33	26.67	34.17	**29.17**	20.83	28.33	27.50	21.77
PRC1-PRC2	26.67	26.67	26.67	37.50	27.50	29.17	32.50	30.00	23.13
PRC1-PRC3	26.67	**30.83**	30.00	34.17	27.50	29.17	29.17	30.00	24.79
PRC2-PRC3	**35.83**	26.67	26.67	35.83	25.00	25.83	23.33	24.17	22.71
PRC1 to PRC3	27.50	30.00	28.33	36.67	26.67	27.50	**36.67**	22.50	23.54
ANOVA+PRC1-to-PRC3	30.83	30.00	**35.83**	**44.17**	26.67	27.50	34.17	27.50	24.06
EEG Features	30.00	26.67	32.50	35.83	25.00	30.00	25.83	27.50	25.31
ANOVA+EEG Features	23.33	**30.83**	35.00	38.33	21.67	30.00	25.00	30.83	23.33

Classification results of ITD-based approaches are given in Tables 4.40-4.47. The effects of selected three PRCs (PRC1, PRC2, and PRC3) and their binary combinations

(PRC1-PRC2, PRC1-PRC3, and PRC2-PRC3) and triple combination (PRC1-to-3) are investigated with 8 different classifiers. Additionally, the effect of ANOVA test-based feature selection is investigated with EEG-based feature set and PRC1-to-3 feature set. In these tables, PRC1, PRC2, or PRC3; indicate that the features for classification are evaluated by utilizing the related PRC. The features are calculated using all three PRCs are indicated as PRC1-to-3. The binary combination features are calculated utilizing; PRC1 and PRC2 is denoted as PRC1-PRC2, PRC1 and PRC3 is denoted as PRC1-PRC3, and PRC2 and PRC3 is denoted as PRC2-PRC3, respectively.

Table 4.41: Finger movement classification performance (%) of ITD based feature sets using the Linear Discriminant Analysis classifier.

Components	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
PRC1	25.83	28.33	33.33	27.50	24.17	30.83	27.50	26.67	**26.25**
PRC2	24.17	28.33	30.00	40.00	24.17	30.83	27.50	27.50	24.79
PRC3	27.50	25.00	32.50	27.50	31.67	**31.67**	21.67	25.83	29.17
PRC1-PRC2	31.67	25.83	35.00	34.17	27.50	26.67	21.67	22.50	29.38
PRC1-PRC3	31.67	26.67	**38.33**	43.33	**.35.83**	30.83	25.00	29.17	29.90
PRC2-PRC3	32.50	25.83	28.33	27.50	25.83	25.83	**32.50**	25.83	28.85
PRC1 to PRC3	31.67	25.00	26.67	33.33	29.17	25.00	20.83	25.00	30.83
ANOVA+PRC1-to- PRC3	**38.33**	**40.00**	37.50	**47.50**	**35.83**	28.33	28.33	**30.00**	**33.54**
EEG Features	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ANOVA+EEG Features	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 4.40 reports the classification results of ITD-based approaches using DT classifier. The best classification result is achieved using ANOVA selected features of PRC1-to-3 set obtained from Subject E (S4) with accuracy of 44.17%. The highest accuracy values are calculated using ITD-based approaches in all subjects. However, in Subject B (S2), the highest accuracy value of 30.83% is achieved in feature sets of PRC1-PRC2 and ANOVA-selected EEG features. Table 4.41 reports the classification results of ITD-based approaches using LDA classifier. The best classification result is achieved using ANOVA selected features of PRC1-to-3 set obtained from Subject E (S4) with accuracy of 47.50%. We aimed to investigate the effect of ITD-based

approaches, but we could not effective comparison with EEG-based approaches. Since, the feature sets of EEG-based approaches are not applicable to LDA classifier.

Table 4.42 reports the classification results of ITD-based approaches using NB classifier. The best classification result is achieved using EEG features obtained from Subject E (S4) with accuracy of 40.83%. We aimed to investigate the effect of ITD-based approaches and EEG-based approaches the results reveal that ITD-based approaches provided the highest classification performances in all subjects except Subject A (S1), Subject E (S4), and subject-independent condition. However, in Subject C (S3), the highest accuracy value of 34.17% is achieved in feature sets of ANOVA-selected PRC1-to-3 and ANOVA-selected EEG features.

Table 4.42: Finger movement classification performance (%) of ITD based feature sets using the Naive Bayes classifier.

Components	**S1 (A)**	**S2 (B)**	**S3 (C)**	**S4 (E)**	**S5 (F)**	**S6 (G)**	**S7 (H)**	**S8 (I)**	**All subjects**
PRC1	22.50	30.83	32.50	35.00	25.00	27.50	30.00	31.67	20.94
PRC2	27.50	25.00	29.17	34.17	24.17	24.17	24.17	25.83	19.38
PRC3	26.67	22.50	25.00	32..50	30.83	25.00	23.33	**32.50**	20.31
PRC1-PRC2	24.17	26.67	29.17	30.83	27.50	22.50	**30.83**	25.00	22.19
PRC1-PRC3	26.67	26.67	33.33	37.50	**38.33**	30.00	29.17	23.33	20.94
PRC2-PRC3	29.17	25.00	30.83	30.00	31.67	25.00	21.67	27.50	21.15
PRC1 to PRC3	23.33	30.00	30.83	35.83	23.33	27.50	23.33	27.50	23.96
ANOVA+PRC1-to -PRC3	30.83	**35.00**	**34.17**	39.17	31.67	**35.83**	**30.83**	30.83	22.81
EEG Features	31.67	30.83	28.33	**40.83**	26.67	22.50	20.83	21.67	**25.42**
ANOVA+EEG Features	**32.50**	25.83	**34.17**	40.00	27.50	29.17	19.17	21.67	23.23

Table 4.42 reports the classification results of ITD-based approaches using NB classifier. The best classification result is achieved using EEG features obtained from Subject E (S4) with accuracy of 40.83%. We aimed to investigate the effect of ITD-based approaches and EEG-based approaches the results reveal that ITD-based approaches provided the highest classification performances in all subjects except Subject A (S1), Subject E (S4), and subject-independent condition. However, in Subject C (S3), the highest accuracy value of 34.17% is achieved in feature sets of ANOVA-selected PRC1-to-3 and ANOVA-selected EEG features.

Table 4.43: Finger movement classification performance (%) of ITD based feature sets using the Support Vector Machine classifier.

Components	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
PRC1	29.17	35.83	40.83	40.00	30.00	38.33	**40.00**	32.50	30.00
PRC2	22.50	25.83	34.17	35.00	30.00	35.83	31.67	24.17	25.73
PRC3	31.67	29.17	30.00	40.00	33.33	29.17	28.33	26.67	27.08
PRC1-PRC2	31.67	31.67	44.17	40.00	24.17	33.33	39.17	27.50	30.52
PRC1-PRC3	35.83	38.33	41.67	47.50	**38.33**	35.00	31.67	29.17	32.19
PRC2-PRC3	29.17	32.50	39.17	37.50	**38.33**	30.83	35.00	32.50	28.13
PRC1 to PRC3	27.50	37.50	45.00	**49.17**	33.33	**38.33**	35.00	35.83	30.63
ANOVA+PRC1 –to-PRC3	**40.00**	**45.00**	**49.17**	**49.17**	35.83	36.67	39.17	**36.67**	**34.48**
EEG Features	30.00	41.67	38.33	45.00	32.50	33.33	29.17	29.17	31.46
ANOVA+EEG Features	27.50	41.67	43.33	47.50	34.17	33.33	30.00	29.17	33.65

Table 4.43 represents the classification results of ITD-based approaches using SVM classifier. The best classification result is achieved using ANOVA-selected PRC1-to-3 features obtained from Subject E (S4) and Subject C (S3) with accuracy of 49.17%. Additionally, the same classification result is achieved using PRC1-to-3 features obtained from Subject E (S4). The experimental results reveal that ITD-based approaches provided the highest classification performances in all subjects.

Table 4.44: Finger movement classification performance (%) of ITD based feature sets using the k-Nearest Neighbours classifier.

Components	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
PRC1	24.17	38.33	35.00	35.83	29.17	34.17	32.50	**37.50**	29.90
PRC2	23.33	25.00	26.67	33.33	31.67	30.00	28.33	23.33	23.02
PRC3	33.33	25.00	34.17	38.33	28.33	30.00	25.83	30.00	26.88
PRC1-PRC2	32.50	30.83	34.17	36.67	26.67	32.50	30.83	30.83	28.54
PRC1-PRC3	31.67	29.17	39.17	**46.67**	32.50	34.17	29.17	30.83	29.48
PRC2-PRC3	32.50	26.67	32.50	35.83	32.50	**35.00**	28.33	26.67	26.25
PRC1 to PRC3	28.33	30.83	35.00	44.17	25.83	31.67	30.83	32.50	26.88
ANOVA+PRC1-to- PRC3	**35.83**	**39.17**	**43.33**	45.83	**35.00**	34.17	**36.67**	35.83	**30.00**
EEG Features	30.00	33.33	33.33	43.33	26.67	29.17	31.67	30.00	27.81
ANOVA+EEG Features	30.00	33.33	40.83	40.00	31.67	31.67	29.17	32.50	28.64

Table 4.44 represents the classification results of ITD-based approaches using k-NN classifier. The best classification result is achieved using PRC1-PRC3 features obtained from Subject E (S4) with accuracy of 46.17%. The experimental results of k-NN classifications reveal that ITD-based approaches provided the highest classification performances in all subjects.

Table 4.45 represents the classification results of ITD-based approaches using EL classifier. The best classification result is achieved using ANOVA-selected PRC1-to-3 features obtained from Subject E (S4) with accuracy of 55.00%. The experimental results of EL classifications show that ITD-based approaches provided the highest classification performances in all subjects.

Table 4.45: Finger movement classification performance (%) of ITD based feature sets using the Ensemble Learning classifier.

Components	**S1 (A)**	**S2 (B)**	**S3 (C)**	**S4 (E)**	**S5 (F)**	**S6 (G)**	**S7 (H)**	**S8 (I)**	**All subjects**
PRC1	29.17	32.50	41.67	35.83	29.17	35.00	37.50	34.17	29.69
PRC2	30.83	30.00	36.67	38.33	28.33	31.67	29.17	30.00	25.10
PRC3	29.17	32.50	34.17	41.67	27.50	32.50	27.50	29.17	26.46
PRC1-PRC2	32.50	34.17	40.00	41.67	33.33	29.17	33.33	32.50	28.85
PRC1-PRC3	35.83	36.67	40.00	43.33	34.17	34.17	38.33	35.83	31.56
PRC2-PRC3	**36.67**	29.17	37.50	45.00	31.67	30.00	28.33	30.83	29.06
PRC1 to PRC3	34.17	35.00	40.83	47.50	32.50	30.83	36.67	31.67	32.08
ANOVA+PRC1-to -PRC3	35.83	**40.83**	**50.83**	**55.00**	**37.50**	**36.70**	**41.67**	**39.17**	**32.60**
EEG Features	30.83	40.00	43.33	39.17	35.83	36.67	26.67	31.67	29.06
ANOVA+EEG Features	29.17	38.33	45.83	39.17	35.00	35.00	27.50	35.00	27.60

Table 4.46 summarizes the classification results of ITD-based approaches using NN classifier. The best classification result is achieved using ANOVA-selected PRC1-to-3 features obtained from Subject C (S3) with accuracy of 53.33%. We aimed to investigate the effect of ITD-based approaches and EEG-based approaches the results reveal that ITD-based approaches provided the highest classification performances in

all subjects except Subject I (S8). However, in Subject G (S6), the highest accuracy value of 38.33% is achieved in feature sets of both ANOVA-selected PRC1-to-3 and ANOVA-selected EEG features.

Table 4.46: Finger movement classification performance (%) of ITD based feature sets using the Neural Networks classifier.

Components	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
PRC1	33.33	29.17	30.83	39.17	23.33	33.33	31.67	29.17	26.25
PRC2	25.00	25.00	31.67	31.67	25.83	25.83	25.83	30.00	24.48
PRC3	32.50	20.83	35.83	35.00	30.00	25.83	32.50	28.33	25.94
PRC1-PRC2	27.50	30.00	43.33	42.50	33.33	30.83	27.50	26.67	28.75
PRC1-PRC3	**34.17**	32.50	40.83	42.50	35.00	30.00	31.67	31.67	30.94
PRC2-PRC3	29.17	29.17	37.50	35.00	34.17	35.83	31.67	30.83	28.96
PRC1 to PRC3	30.00	33.33	45.83	**48.33**	**37.50**	32.50	30.00	34.17	29.27
ANOVA+PRC1 to PRC3	**34.17**	**42.50**	**53.33**	45.83	**37.50**	**38.33**	**35.00**	31.67	**31.88**
EEG Features	28.33	35.83	42.50	39.17	35.00	29.17	26.67	32.50	28.96
ANOVA+EEG Features	25.83	35.00	41.67	42.50	36.67	**38.33**	23.33	**35.00**	30.42

Table 4.47: Finger movement classification performance (%) of ITD based feature sets using the Kernel Approximation classifier.

Components	S1 (A)	S2 (B)	S3 (C)	S4 (E)	S5 (F)	S6 (G)	S7 (H)	S8 (I)	All subjects
PRC1	20.00	**25.00**	25.83	24.17	26.67	25.83	**30.83**	20.83	23.33
PRC2	26.67	**25.00**	23.33	30.00	23.33	17.50	23.33	24.17	19.27
PRC3	27.50	20.00	27.50	35.83	21.67	22.50	21.67	22.50	21.88
PRC1-PRC2	22.50	21.67	20.83	22.50	19.17	19.17	26.67	25.00	19.58
PRC1-PRC3	25.83	24.17	27.50	39.17	**27.50**	22.50	29.17	20.00	24.48
PRC2-PRC3	25.00	20.83	27.50	33.33	20.00	23.33	21.67	19.17	24.27
PRC1 to PRC3	24.17	24.17	27.50	**40.83**	15.00	25.83	25.00	25.83	23.33
ANOVA+PRC1 –to-PRC3	21.67	18.33	26.67	31.67	19.17	22.50	22.50	25.83	21.88
EEG Features	25.83	**25.00**	38.33	32.50	25.00	**30.00**	27.50	**29.17**	24.17
ANOVA+EEG Features	**29.17**	22.50	**34.17**	36.67	**27.50**	26.67	20.00	26.67	**25.31**

Table 4.47 summarizes the classification results of ITD-based approaches using KA classifier. The best classification result is achieved using PRC1-to-3 features obtained from Subject E (S4) with accuracy of 40.83%. We aimed to investigate the effect of ITD-based approaches and EEG-based approaches, the results reveal that EEG-based approaches provided the highest classification performances in all subjects except Subject E (S4) and Subject H (S7). On the hand, the same highest accuracy values (25.00% and 27.50%) are calculated using both ITD-based and EEG-based approaches for Subject B (S2) and Subject F (S5).

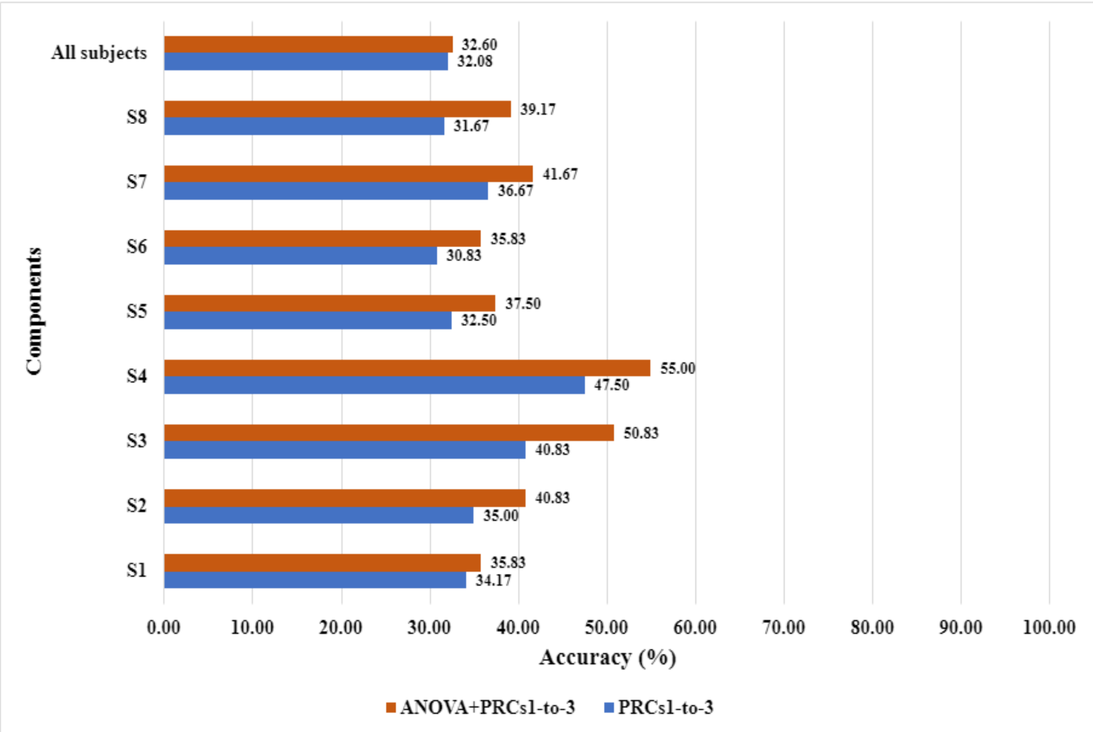


Figure 4.3: The comparision of the effects of PRC1-to-3 and ANOVA-selected PRC1-to-3 based on the EL classification.

ITD-based approaches revealed that the best classification results are achieved using ITD-based features (especially in ANOVA-selected PRC1-to-3) features in most of the classifiers except NB and KA classifiers. Therefore, our presented different combinations of PRCs are improved the classifiers performance. Additionaly, when we compared the effectiveness of PRC1-to-3 and ANOVA-selected PRC1-to-3, we observed that ANOVA-selected features obtained better results and improved classifier performances. The highest accuracy value of ITD-based approaches is achieved using ANOVA-selected PRC1-to-3 features and EL classifier as shown in Figure 4.3. On the other hand, the highest classification performances are calculated using ITD-based features obtained from Subject E (S4). Therefore, the experimental noted that ITD is an effective time-frequency representation model for classification of finger movement and provides better results than our first time-frequency representation model (WT). It can be used as an effective feature extraction method to analyze time-frequency domain of EEG signals in different EEG-based analysis.

The selected statistically significant ITD-based time-frequency distribution was examined in 19 EEG channels. The list of ITD-based time-frequency features with their abbreviations is available in Table 4.48. Channel-based ANOVA-selected statistically significant ITD-based time-frequency domain feature distribution for subject-dependent finger movement classification is given in Table 4.49. Among 30 different ITD-based features, some features such as sample entropy, Hjorth parameters (Mobility), and Hjorth parameters (Complexity) for PRC1, PRC2, and PRC3 were mostly indicated and selected as statistically significant features on almost all channels and 8 subjects. When the effect of the channels on the selection of statistically significant features was examined, it was observed that there was a balanced distribution and a large number of significant features were selected in many subjects from all channels. As a result, feature selection from all channels and certain features with ANOVA improved the classification performance in all classifiers except NB and KA.

Table 4.48: ITD-based time-frequency domain features.

ITD-based time-frequency domain features			
I_1	Power for PRC1	I_{16}	Hjorth parameters (Complexity) for PRC2
I_2	Mean for PRC1	I_{17}	First higher order moment for PRC2
I_3	Sample entropy for PRC1	I_{18}	Second higher order moment for PRC2
I_4	Hjorth parameters (Activity) for PRC1	I_{19}	Third higher order moment for PRC2
I_5	Hjorth parameters (Mobility) for PRC1	I_{20}	Fourth higher order moment for PRC2
I_6	Hjorth parameters (Complexity) for PRC1	I_{21}	Power for PRC3
I_7	First higher order moment for PRC1	I_{22}	Mean for PRC3
I_8	Second higher order moment for PRC1	I_{23}	Sample entropy for PRC3
I_9	Third higher order moment for PRC1	I_{24}	Hjorth parameters (Activity) for PRC3
I_{10}	Fourth higher order moment for PRC1	I_{25}	Hjorth parameters (Mobility) for PRC3
I_{11}	Power for PRC2	I_{26}	Hjorth parameters (Complexity) for PRC3
I_{12}	Mean for PRC2	I_{27}	First higher order moment for PRC3
I_{13}	Sample entropy for PRC2	I_{28}	Second higher order moment for PRC3
I_{14}	Hjorth parameters (Activity) for PRC2	I_{29}	Third higher order moment for PRC3
I_{15}	Hjorth parameters (Mobility) for PRC2	I_{30}	Fourth higher order moment for PRC3

Table 4.49: Channel-based ANOVA-selected statistically significant feature distribution for subject-dependent finger movement classification in ITD-based time-frequency domain feature set.

F_{id}	Channels																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**T**
I_1	3	2		1	2	2		2	3	3	2	1	4	2	1	2	3	2	3	38
I_2	2	2	1	2		2	1	2	4	3	4	3	4	3	1	3	3	1	1	42
I_3	5	6	3	3	2	4	2	5	4	5	6	6	5	5	3	4	5	4	4	81
I_4	3	2		1	2	2		2	3	3	2	1	4	1	2	2	3	2	3	38
I_5	6	7	6	6	6	7	6	6	5	5	4	7	7	7	6	8	6	6	5	116
I_6	7	6	6	6	6	5	6	6	4	6	7	6	5	4	4	7	4	6	5	106
I_7		1	1	1	1	1	1	1		1	1	3	3	2			1	2	1	21
I_8		1	1	1	1	1	1	1		1	1	3	3	2			1	2	1	21
I_9		1	1	1	1	1	1	1		1	1	3	3	2			1	2	1	21
I_{10}		1	1	1	1	1	1	1		1	1	3	3	2			1	2	1	21
I_{11}	2	4	3	4	1	1	2	3	3	4	3	4	6	4	2	3	4	2	1	56
I_{12}	3	2			1	1	1		1				1	1	1	1				13
I_{13}	5	5	6	6	5	5	5	5	6	4	7	6	4	8	5	3	7	5	4	101
I_{14}	2	4	3	4	1	1	2	3	3	4	3	4	6	4	2	4	4	2	1	57
I_{15}	6	5	6	5	5	5	5	6	7	5	6	6	7	8	4	5	6	7	7	111
I_{16}	5	3	6	6	5	5	5	5	5	5	6	7	4	8	5	4	4	6	6	100
I_{17}	1	3	2		1	3	4	2	1	3	1	2	3	5	1	3	3	3	3	44
I_{18}	1	3	2		1	3	4	2	1	3	1	2	3	5	1	3	3	3	3	44
I_{19}	1	3	2		1	3	4	2	1	3	1	2	3	5	1	3	3	3	3	44
I_{20}	1	3	2		1	3	4	2	1	3	1	2	3	5	1	3	3	3	3	44
I_{21}	7	7	2	5			2	4	2	3	1	3	4	3	2	4	6	2	2	59
I_{22}	6	5	1	2			1		2	1	1	4			1	1	3		1	29
I_{23}	7	6	5	4	4	5	3	5	4	3	6	7	6	4	5	4	6	6	3	93
I_{24}	7	7	1	5			2	5	2	2		4	3	3	2	4	5	2	2	56
I_{25}	7	6	5	4	4	4	3	7	5	3	5	7	2	5	6	4	6	7	3	93
I_{26}	7	7	4	4	5	6	4	6	3	2	3	5	5	5	6	5	6	7	4	94
I_{27}	7	7	2	4	4	2		2			1	4	2	2		1	6	2		46
I_{28}	7	7	2	4	4	2		2			1	4	2	2		1	6	2		46
I_{29}	7	7	2	4	4	2		2			1	4	2	2		1	6	2		46
I_{30}	7	7	2	4	4	2		2			1	4	2	2		1	6	2		46
T	1-22	1-30	78	88	73	79	70	92	70	77	78	1-17	1-09	1-11	62	84	1-21	95	71	1727

Channel-based ANOVA-selected statistically significant ITD-based time-frequency domain feature distribution for subject-independent finger movement classification is given in Table 4.49. Among 30 different ITD-based features, some features such as sample entropy for PRC1, PRC2, and PRC3, Hjorth parameters (Mobility) for PRC2 and PRC3, and Hjorth parameters (Complexity) for PRC1, PRC2, and PRC3 were mostly indicated and selected as statistically significant features on almost all channels. When the effect of the channels on the selection of statistically significant features was examined, it was observed that there was a balanced distribution and a large number of significant features were selected from all channels. Therefore, feature selection from all channels and certain features with ANOVA improved the classification performance in all classifiers except NB and KA.

Table 4.50: Channel-based ANOVA-selected statistically significant feature distribution for subject-independent finger movement classification in ITD-based time-frequency domain feature set.

F_{id}	Channels																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	T
I_1																				0
I_2	■					■						■	■	■	■					6
I_3		■				■			■	■		■	■	■	■	■		■	■	11
I_4																				0
I_5		■				■						■		■	■					5
I_6		■	■	■	■	■					■	■		■			■	■		10
I_7											■									1
I_8											■									1
I_9											■									1
I_{10}											■									1
I_{11}		■						■	■	■	■			■						6
I_{12}	■	■							■									■		4
I_{13}	■	■	■	■	■	■	■	■	■	■	■	■	■			■	■	■		16
I_{14}		■						■	■	■	■			■						6
I_{15}	■	■		■	■	■	■	■	■	■	■	■	■			■	■	■		15
I_{16}	■	■		■	■	■	■	■	■	■	■	■	■	■		■	■	■		16
I_{17}		■					■					■	■	■						5
I_{18}		■					■					■	■	■						5
I_{19}		■					■					■	■	■						5
I_{20}		■					■					■	■	■						5
I_{21}	■	■	■	■								■	■	■			■			8
I_{22}	■	■	■	■								■				■	■			7
I_{23}	■	■		■								■	■				■			6
I_{24}	■	■	■	■								■	■	■			■			8
I_{25}	■	■										■	■				■			5
I_{26}	■	■	■	■	■	■	■	■				■	■	■	■	■	■	■	■	16
I_{27}	■	■	■	■	■							■	■	■			■			9
I_{28}	■	■	■	■	■							■	■	■			■			9
I_{29}	■	■	■	■	■							■	■	■			■			9
I_{30}	■	■	■	■	■							■	■	■			■			9
T	15	23	10	13	9	8	8	6	7	6	10	21	18	18	4	6	14	7	2	205

In our performed finger movement classification analyzes, we investigated the effects of different feature extraction approaches and PCA-based and statistically significance-based feature selection methods. We performed time-domain, frequency-domain, WT-based time-frequency domain, and non-linear feature sets. In addition to these feature set, we investigated two different combinations of these features to improve classifier performance. A total of six feature set are created and analyzed separately. We presented four different approaches using PCA and ANOVA test for each feature set to reveal effectiveness of ANOVA test. These approaches are (i) all features, (ii) PCA-selected features, (iii) ANOVA-selected features, and (iv) ANOVA and PCA selected features from corresponding feature set. The presented approaches are analyzed using various classifiers, separately. Hence the effects of different feature sets and feature selection methods are analyzed for finger movement classification. In addition to our WT-based time-frequency analysis, we performed another time-frequency representation (ITD) model to compare their effectiveness for classification finger movement. In this approach we investigated effects of the selected PRCs (PRC1, PRC2, and PRC3) and their different combinations. We also used ANOVA test to improve classifier performance. The obtained ITD-based feature sets are classified using different machine learning algorithms. In order to demonstrate the improvements of utilizing ITD approaches, the same features were calculated from the EEG signal itself, and classification step is repeated.

We applied the proposed 5 different feature extraction approaches for classification of NoMT condition and five finger movements MI tasks of 19-channel EEG signals after obtaining of 1 sec MI EEG segments. A total of 24 time-domain, 15 frequency-domain, 15 WT-based time-frequency domain, 10 different and 4 non-linear features are evluated from each EEG segment. These time-domain, frequency-domain, WT-based time-frequency domain, ITD-based time-frequency domain, and non-linear feature sets, their two different combination feature sets, and their features selected feature sets of all feature sets were classified utilizing DT, DA, NB, SVM, k-NN, EL, NN, and KA. The performances of different feature sets and the effectiveness of PCA and ANOVA were investigated and compared.

Among all feature sets, performance of WT-based time-frequency feature set was observed to be poor for finger movement classification and the performance of combination feature sets was found to be higher especially for SVM and EL classifiers. The highest accuracy value (59.17%) is achieved using ANOVA-selected TD+FD+WT+P and SVM algorithm. We observed that our Poincare measures-based non-linaer features are improved the classifier performance when these features included the first combination feature set including time-domain, frequency-domain, and WT-based time-frequency features. The highest accuracy values in different feature sets are generally obtained using SVM and EL algorithms. Additionally, we investigated the effect of ANOVA-based feature selection and observed that the classification performances of ANOVA-selected features are generally improve the classification performance determining discriminative features for finger movement classification.

In addition, we investigated ITD-based time-frequency features for finger movement classification. Experimental results performed with different ITD-based feature sets revealed that the combinations of the different PRCs improve classifier performance and the highest classification performances are obtained using ANOVA-selected PRC1-to-3 features. ANOVA test-based feature selection process helps improve the classification performance. To reveal the effect of ITD approaches, when we compared ITD-based approaches and EEG-based approaches, we observed that highest accuracy values are mostly achieved using ITD-based approaches in all classifiers except NB and KA classifiers. On the other hand, we compared our two time-frequency representation models (WT and ITD), the classification performances of ITD-based approaches are higher than WT-based approaches. Therefore, we noted that ITD algorithm is an effective time-frequency representation model and carries the most useful information than WT-based approaches for classification of finger movement. Thus, the encouraging results of ITD-based feature extraction, the combination of different feature extraction methods and statistically significance-based feature selection showed that the proposed approaches may be used for EEG-based analysis.

The performance of the proposed methods for finger movement classification is compared with the success of the finger movement classification studies performed using same data set in literature (given in Table 4.51). In some studies [22, 66], subject-independent analysis was performed. When these studies are examined, higher performance values are reported than our proposed methods. However, in [22], using

Table 4.51: Performance comparison of finger movement classification studies.

Ref.	Subject condition	Number of channels	Number of classes	Feature	Classifier	Accuracy (%)
[22]	SI/8	19	5	EEG time series Noise addition Sliding window	CNN	57.50
[71]	SD/4	19	5	CSP	Random forest	51.00-56.00
[72]	SD/4	19	5	CSP	Autonomous deep learning	74.61-77.75
[73]	SD/8	19	5	Multi-class CSP Complex fourier amplitudes	SVM	23.90-58.30
[66]	SD/8	19	5	EEG sunbbands power Forier transform amplitudes EEG time series	SVM	20.00-60.00
[66]	SI/8	19	5	EEG sunbbands power Forier transform amplitudes EEG time series	SVM	43.00
[69]	SD/8	4	5	Spectrogram features	SVM	21.20-66.60
[67]	SD/4	4	5	EEG time series	Deep learning	80.10-91.70
[70]	SD/8	4	5	EMD	BiLSTM	66.00-76.13
This study	**SI/8**	**19**	**6**	**TD**	**SVM**	**36.20**
				FD+ANOVA	**SVM**	**30.45**
				WT	**EL**	**26.60**
				P+ANOVA	**SVM**	**31.79**
				TD+FD+WT+ANOVA	**SVM**	**38.70**
				TD+FD+WT+P+ANOVA	**SVM**	**39.30**
				ITD+ANOVA	**SVM**	**34.48**
This study	**SD/8**	**19**	**6**	**TD+ANOVA**	**SVM**	**33.30-57.50**
				FD+ANOVA	**EL**	**29.17-55.00**
				WT+ANOVA	**SVM**	**25.83-36.67**
				P+ANOVA	**SVM**	**29.17-50.00**
				TD+FD+WT+ANOVA	**SVM**	**35.83-55.83**
				TD+FD+WT+P+ANOVA	**SVM**	**32.50-59.17**
				ITD+ANOVA	**EL**	**35.83-55.00**

of CNN architecture increased complexity of classification process. In other studies [66, 67, 69, 70-73], subject-dependent analysis was performed. In some of these studies [67, 69, 70], the channel reduction is performed and only 4 out of 19 EEG channels selected. Among these studies, in [67, 70], higher performance values are presented than our proposed study. In [67], in addition to channel reduction, EEG data of only 4 subjects was used and their approaches inludes higher computational complexity due to training time of deep learning approach. In another study [70], BiLSTM structure which increased classifier complexity was used for classification. In [66], all channels and all subjects were used as our analysis for their subject-dependent study and the highest accuracy value of 60.00% was achieved, but it has also been observed that accuracy value of 20.00% was achieved in some subjects. In [71, 73] all EEG channels are used as our analysis, but the classification results of our methods are higher than these studies. In fact, in [71], EEG data of only 4 subjects were used. In another study [72], subject-dependent analysis was performed using deep learning structure and EEG data of only 4 subjects. Their experimental results were higher than our experimental results. However, their analysis includes higher complexity than our proposed methods. Our studies include low computational complexity in terms of feature extraction, feature selection, and classification methods. In addition, our statistically significance feature distribution examinations in different feature sets revealed that the statistically significant feature density selected from the channels and the selected feature types may vary in different feature sets. Therefore, in each study, significant and relevant feature types and channels can be determined by first extracting features from all channels and using ANOVA-based feature selection, and the study can continue with these channels and features. These analyzes are included in detail in our studies. Since our studies did not focus on the same channels in all feature sets, feature extraction, feature selection and classification processes were continued by using the information of all channels. Using different approaches, the promising classification results were achieved in finger movement classification studies.

Chapter 5

Conclusion

The accurate decoding of MI tasks plays an important role in BCI design in order to improve ADL of indivuduals who exposure motor impairment. There are different neuroimaging methods to provide electrophysiologic activity of the brain for BCI studies. Despite various brain imaging modalities, the EEG signal based BCI system design has mostly performed due to its low cost and ease of recording, high temporal resolution. However, EEG signals have non-linear and non-stationary characteristic structure and these drawbacks make MI EEG classifications is quite difficult task. Hence, the various methods have been introduced to accurate decoding of MI EEG signals with high-quality processing of EEG signals in the literature. This thesis aimed to propose various feature extraction methods with different feature selection methods and machine learning algorithms by using EEG signals of two different MI task classification, extremity movement and finger movement.

Firstly, the extremity movement task classification approaches are presented. In this thesis, four different feature extraction methods that can be utilized in the classification of binary-class (right hand and left hand) and four-class (right hand, left hand, both feet, and tongue) extremity movements have been introduced with the combination of 22-channel EEG signals and machine learning algorithms. An open-available BCI Competition IV-IIa dataset was used for EEG signal analysis in the extremity movement task classification studies. Before feature extraction step, 3 sec MI EEG segments are decomposed from EEG signals which belongs to four different class categories. Then, 24 diferent time-domain, 15 FFT-based frequency domain, 15 WT-based time-frequency, and 4 Poincare plot-based non-linear features are calculated for each EEG segments. This process is performed for all 22 EEG channels. Four different feature sets including time-domain, frequency-domain, time-frequency domain, and non-linear are provided. In addition to these sets, two combination feature sets of different features are created and analyzed for investigation of effects of different feature sets. The first combination feature set consists of time-domain, frequency-domain, and time-frequency domain features, while the second one includes all features of first combination with addition of non-linear features. Additionaly, the statistically significance-based feature selection methods such as the independent t-test and ANOVA test are used to improve classifier performance selecting relative and discriminative MI EEG features for binary and multiple extremity movement task classifications, respectively. The effect of ANOVA and the independent t-test is investigated separately in each feature set. For four-class extremity movement task classification, the six obtained feature sets are classified by a total of 30 different classification processes according to the eight basic classifiers, while for binary classifications, a total of 31 classification processes are performed with the addition of LR-based classification process.

In our simulations for the binary-class extremity movement task classification, the highest classification accuracy value was obtained by using non-linear feature set approach where the relavent information about MI tasks may be supplied more clearly. When we analyzed the effect of all feature sets, performance of time-frequency feature set was observed to be poor for binary-class extremity movement task classification and the performance of non-linear feature sets was found to be higher especially for all classifiers except NB. Additionally, among two different combination sets, the highest value was achieved with the first combination set. Note that, working with the independent t-test as feature selection method, generally improved the classifier performance in all feature sets except time-domain feature set. When we observed the effect of 9 classifier algorithms, DA algorithm achieved the highest accuracy of binary-class extremity movement task classification with using non-linear feature set. However, the highest accuracy value in different feature sets is generally obtained by EL algorithm-based classifications (shown in Figure 3.3a and Figure 3.3b).

In our simulations for the multi-class extremity movement task classification, the highest classification accuracy value was obtained by the second feature set (TD+FD+WT+P+ANOVA) approach where the relavent information about MI tasks may be supplied more clearly. Among four feature sets (time-domain, frequency-domain, time-frequency domain, and non-linear), performance of time-frequency feature set was observed to be poor for four-class extremity movement task classification the performance of non-linear feature sets was found to be higher especially for SVM and EL classifiers. The most successful non-linear feature set is “All lags” non-linear feature set including all non-linear features which are calculated for 10 different lag conditions. It was shown that, despite the high classification performance of the SVM algorithm with the non-linear feature set among the four feature sets, it did not perform the same better performance with the combination feature set. In addition, we observed that the highest performance in the proposed approaches was generally achieved with the EL algorithm in the classifications performed on all features sets (time-domain, frequency-domain, time-frequency domain, and combination sets) except the non-linear feature sets (as shown in Figure 3.4a and Figure 3.4b). The experimental results performed using ANOVA-selected feature sets revealed that ANOVA is improved the classifier performance in all proposed feature sets determining discriminative and informative MI EEG features from the corresponding feature sets.

In the second section of thesis, various feature extraction approaches and two different feature selection methods such as ANOVA and PCA have been presented to classify the EEG signals of finger movements. The subject-dependent and subject-independent finger movement classification analyzes are performed. An open-available large electroencephalographic MI dataset was used for EEG signal analysis in the finger movement classification studies. NoMT and 5F 19-channel EEG signals are used for our analysis. 100 sample determined for each class category and a total of 600 sample used for feature extraction step to provide balanced data distribution. As our previous feature extraction step performed for the extremity movement task classification, the same six feature sets are extracted from EEG signals of finger movements and NoMT condition. In feature selection, PCA-based feature selection is added to process to reveal effect of ANOVA test. In this direction, four different feature sets generated based on different feature selection methods, ANOVA and PCA, from our six feature sets. These generated features sets are (i) all feature set, (ii) PCA-selected feature set, (iii) ANOVA-selected feature set, and ANOVA and PCA-selected feature set from the corresponding feature set (TD, FD, WT, P, TD+FD+WT, and TD+FD+WT+P). These feature sets are evaluated utilizing 30 different classification processes according to the eight basic classifiers. In addition to all feature sets, ITD-based approach is used to obtain time-frequency features as alternative to WT-based time-frequency features. In this process, PRCs are extracted from each EEG segments using ITD. To defining of informative PRC is important step in this process. We used energy-based examinations to define PRCs and selected first three PRCs for our analysis due to their high-frequency content. 10 different ITD-based time-frequency features are evaluated from the corresponding PRCs. In addition to first three PRCs (PRC1, PRC2, and PRC3), different binary-combinations of PRCs (PRC1-PRC2, PRC1-PRC3, and PRC2-PRC3) and their triple combination (PRC1-to-3). To reveal the effect of ITD, the same features from EEG segments itself without application of ITD. Additionaly, ANOVA-based feature selection is performed for PRC1-to-3 and EEG-based feature sets to improve classifier performance. A total of 8 ITD-based feature sets and 2 EEG-based feature sets are classified with 30 different classification processes under eight classifiers algorithms.

The simulation results revealed that WT-based time-frequency feature set-based classifications obtained worse accuracy values and the combination feature set (TD+FD+WT+P) achieved better results for finger movement classification. The highest accuracy value (59.17%) is calculated utilizing ANOVA-selected TD+FD+WT+P feature set and SVM algorithm. The results noted that our Poincare measures-based non-linaer features are improved the classifier performance when these features included the first combination feature set (TD+FD+WT) including time-domain, frequency-domain, and WT-based time-frequency features. On the other hand, the results show that ANOVA-selected statistically significant features are generally improve the classification performance obtaining informative features in all feature sets for finger movement classification. When we analyzed the simulation results of ITD-based approaches, we observed that the combinations of the different PRCs obtain better results and the highest classification performances are achieved utilizing ANOVA-selected PRC1-to-3 features. ANOVA-based feature selection generally improves the classifier performance in both PRC1-to-3 feature set and EEG-based feature set. We compared the ITD-based and EEG-based approaches, the results noted that highest accuracy values are mostly calculated utilizing ITD-based approaches in all classification algorithms except NB and KA classifiers. Although WT-based approaches have been mostly used in classification of MI signals in the literature, in our study, ITD-based time-frequency approaches provide better results than WT-based time-frequency approaches for finger movement classification. Hence, the results demonstrate that ITD-based approaches, which obtain the time-frequency representation of EEG signals, can be used successfully in discrimination of finger movements.

The objectives, which are provided in our thesis, are summarized as follows;

- Various feature extraction approaches such as time-domain, frequency-domain, WT-based time-frequency domain, ITD-based time-frequency domain, non-linear features and their different combinations are investigated and the advantages this Poincare plot-based non-linear features and the combinations sets are presented in deatil for binary and multiple extremity movement task classification and finger movement classification.

- In addition to the literature studies which has been mostly performed using spatial features such as CSP and its different versions, Poincare plot measures-based non-linear feature extraction has been proposed and the effect of different lag values are investigated. It has been shown that the successful classification results were achieved using Poincare plot measures.

- It has been noted that the ITD method can be utilized successfully in the classification of finger movement.

- Different combinations of PRCs extracted using ITD approach are investigated for the first time with various classifier algorithms for the classification of MI tasks and successful performance evaluation results are achieved.

- The simulation results performed with the statistically significance-based feature selection methods (ANOVA test and the independent t-test) reveal that they improved the classifier performance selecting relative and discriminative features and it can be used as feature selection method in EEG-based MI task classification.

- The channel-based distributions of statistically significant features determined by the statistically significance-based feature selection method were examined for both extremity movement task and finger movement classifications. The results revealed that the statistically significant feature density selected from the channels and the selected feature types may vary in different feature sets. Therefore, instead of eliminating the channels at first, it was concluded that extracting features from all channels and then determining significant and relevant EEG channels and features by statistically significance-based feature selection method is an effective way.

- In the literature, the performances of the machine learning algorithms that were not analyzed before were calculated in terms of MI task classification and the most successful classification algorithm was found to be the SVM and EL lassifiers.

- In the literature, passive mode (NoMT condition) has ignored for finger movement classification studies, we presented a six-class finger movement classification study including of EEG signals which belong to NoMT class to provide a more realistic BCI design for indivuduals, who suffering from motor disabilities to the author's best knowledge.

References

[1] Nicolas-Alanso LF, Gomez-Gil J. Brain computer interfaces, a review. Sensors 2012; 12(2): 1211-1279.

[2] He B, Yuan H, Meng J, Gao S. Brain-computer interfaces. Neural Engineering 2020; 131-183.

[3] Russell S, Norvig P. Artificial intelligence—a modern approach (3rd edition): Prentice Hall; 2018.

[4] Singh AK, Wang YK, King JT, Lin CT. Extended interaction with a BCI video game changes resting-state brain activity. IEEE Transactions on Cognitive and Developmental Systems 2020; 12(4): 809-823.

[5] Spatro R, Chella A, Allison B, Giardina M, Sorbello R, Tramonte S, *et al*. Reaching and grasping a glass of water by locked-in ALS patients through a BCI-controlled humanoid robot. Frontiers in Human Neuroscience 2017; 68.

[6] Saha S, Mamun KA, Ahmed K, Mostafa R, Naik GR, Darvishi S, *et al*. Progress in brain computer interface: Challenges and opportinities. Frontiers in Systems Neuroscience 2021; 15: 578875.

[7] Carlson T, Del J, Millan R. Brain-controlled wheelchairs: A robotic architecture. IEEE Robotics & Automation Magazine 2013; 20(1): 65-73.

[8] Meng J, Zhang S, Bekyo A, Olsoe J, Baxter B, He B. Noninvasive electroencephalogram based control of a robotic arm for reach and grasp tasks. Scientific Reports 2016; 6(1): 38565.

[9] Bouton CE, Shaikhouni A, Annetta NV, Bockbrader MA, Friedenberg DA, Nielson DM, *et al*. Restoring cortical control of functional movement in a human with quadriplegia. Nature 2016; 533(7602): 247.

[10] DeBettencourt MT, Cohen JD, Lee RF, Norman KA, Turk-Browne NB. Closed-loop training of attention with real-time brain imaging. Nature Neuroscience 2015; 18(3): 470-478.

[11] Penaloza CI, Nishio S. BMI control of a third arm for multitasking. Science Robotics 2018; 3(20): eaat1228.

[12] Biasiucci A, Leeb R, Iturrate I, Perdikis S, Al-Khodairy A, Corbet T, *et al*. Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke. Nature Communications 2018; 9(1): 2421.

[13] Mason SG, Birch GE. A general framework for brain-computer interface design. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2003; 11(1): 70-85.

[14] Waziri A, Claassen AJ, Stuart RM, Arif H, Schmidt JM, Mayer SA, Badjatia N, *et al*. Intracortical electroencephalography in acute brain injury. Annals of Neurology 2009; 66(3): 366-77.

[15] Chao ZC, Nagasaka Y, Fujii N. Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkeys. Frontiers in Neuroengineering 2010; 3.

[16] Bandettini PA, Wong EC, Hinks RS, Tikofsky RS, Hyde JS, Time course EPI of human brain function during task activation. Magnetic Resonance in Medicine 1992; 25(2): 390-397.

[17] Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, *et al*. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proceeding of the National Academy of Sciences 1992; 89(12), 5675-5679.

[18] Shirley C, Ward T, Markham C, McDarby G. On the suitability of near-infrared (NIR) systems for next-generation brain-computer interfaces. Physiological Measurement. 2004; 25: 815.

[19] Sitaram R, Zhang H, Guan C, Thulasidas M, Hoshi Y, Ishikawa A, *et al*. Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain–computer interface. Neuroimage 2007; 34: 1416-1427.

[20] Baillet S, Mosher JC, Leahy RM. Electromagnetic brain mapping. IEEE Signal Processing Magazine 2001; 18: 14-30.

[21] Kauhanen L, Nykopp T, Lehtonen J, Jylanki P, Heikkonen J, Rantanen P, *et al*. EEG and MEG brain-computer interface for tetraplegic 122 patients. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2006; 14: 190-193.

[22] Zahra A, Kanwal N, ur Rehman N, Ehsan S, McDonald-Maier KD. Seizure detection from EEG signals using multivariate empirical mode decomposition. Computers in Biology and Medicine 2017; 88: 132-141.

[23] Bajaj V, Pachori RB. Classification of seizure and nonseizure EEG signals using empirical mode decomposition. IEEE Transactions on Information Technology in Biomedicine 2011; 16(6): 1135-1142.

[24] Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain–computer interfaces for communication and control. Clinical Neurophysiology 2002; 113(6): 767-791.

[25] Ajiboye AB, Willett FR, Young DR, Memberg WD, Murphy BA, Miller JP, *et al*. Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: A proof-of-concept demonstration. The Lancet 2017; 389(10081): 1821-1830.

[26] Wolpaw JR, Boulay CB. Brain-computer interfaces: Springer; 2010.

[27] Ramadan RA, Vasilakos AV. Brain computer interface: control signals review. Neurocomputing 2017; 223: 26-44.

[28] Li Y, Wen PP. Modified CC-LR algorithm with three diverse feature sets for motor imagery tasks classification in EEG based brain-computer interface. Computer Methods and Programs in Biomedicine 2014; 113(3): 767-780.

[29] Jeannerod, M. Mental imagery in the motor context. Neuropsychologia 1995; 33: 1419-1432.

[30] Altaheri H, Muhammed G, Alsulaiman M, Amin SU, Altuwaijiri GA, Abdul W, *et al*. Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (MI) signals: A review. Neural Computing and Applications 2023; 35(20): 14681-14722.

[31] Baali H, Khorshidtalab A, Mesbah M, Salami MJ. A transform-based feature extraction approach for motor imagery tasks classification. IEEE Journal of Translational Engineering in Health and Medicine 2015; 3: 1-8.

[32] Tiwari S, Goel S, Bhardwaj A. MIDNN-a classification approach for the EEG based motor imagery tasks using deep neural network. Applied Intelligence 2022; 52(5): 4854-4843.

[33] Musallam YK, Alfassam NI, Muhammad G, Amin SU, Alsulaiman M, Abdul W, *et al*. Electroencephalography-based motor imagery classification using

temporal convolutional network fusion. Biomedical Signal Processing and Control 2021; 69: 102826.

[34] Lotte F, Congedo M, Lecuyer A, Lamrche F, Arnaldi B. A review of classification algorithms for EEG-based brain-computer interfaces. Journal of Neural Engineering 2007; 4(2): R1.

[35] Saa JFD, Cetin M. Discriminative methods for classification of asynchronous imaginary motor tasks from EEG data. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2013; 21(5): 716-724.

[36] Ha KW, Jeong JW. Motor imagery EEG classification using capsule networks. Sensors 2019; 19(13): 2854.

[37] Garcia-Laencina PJ, Rodriguez-Bermudez G, Roca-Dorda J. Exploring dimensionality reduction of EEG features in motor imagery task classification. Expert Systems with Applications 2014; 41(11): 5285-5295.

[38] Luo J, Gao X, Zhu X, Wang B, Lu N, Wang J. Motor imagery EEG classification based on ensemble support vector learning. Computer Methods and Programs in Biomedicine 2020; 193: 105464.

[39] Pawar D, Dhage S. Feature Extraction Methods for Electroencephalography based Brain-Computer Interface: A Review. International Journal of Computer Science (IAENG) 2020; 47(3): 501-515.

[40] Degirmenci M, Yuce YK, Isler Y. Motor imaginary task classification using statistically significant time-domain EEG features. In 2022 30th Signal Processing and Communications Applications Conference (SIU). IEEE; 2022. 1-4.

[41] Djamal EC, Abdullah MY, Renaldi F. Brain computer interface game controlling using fast Fourier transform and learning vector quantization. Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 2017; 9(2-5): 71-74.

[42] Oppenheim AV. Discrete-time signal processing. Pearson Education India; 1999.

[43] Kevric J, Subasi A. Comparision of signal decomposition methods in classification of EEG signals for motor-imagery BCI system. Biomedical Signal Processing and Control 2017; 31: 398-406.

[44] Chaudhary S, Taran S, Bajaj V, Siuly S. A flexible analytic wavelet transform based approach for motor-imagery tasks classification in BCI applications. Computer Methods and Programs in Biomedicine 2020; 187: 105325.

[45] Verma NK, Rao LVS, Sharma SK. Motor imagery EEG signal classification on DWT and crosscorrelated signal features. In 2014 9th International Conference on Industrial and Information Systems (ICIIS). IEEE; 2014. 1-6.

[46] Mohamed EA, Yusoff MZ, Adam IK, Hamid EA, Al-Shargie F, Muzammel M. Enhancing EEG signals in brain computer interface using intrinsic time-scale decomposition. In Journal of Physics: Conference Series 2018; 1123(1): 012004.

[47] Alam ME, Samanta B. Empirical mode decomposition of EEG signals for brain computer interface. In 2017 SoutheastCon. IEEE; 2017. 1-6.

[48] Wu W, Chen Z, Gao X, Li Y, Brown EN, Gao S. Probabilistic common spatial patterns for multichannel EEG analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 2014; 37(3): 639-653.

[49] Ang KK, Chin ZY, Zhang H, Guan C. Filter bank common spatial pattern (FBCSP) in brain-computer interface. In 2018 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Inte). IEEE; 2018. 2390-2397.

[50] Kumar S, Sharma A, Mamun K, Tsunoda T. A deep learning approach for motor imagery EEG signal classification. In 2016 3rd Asi-Pacific World Congress on Computer Science and Engineering (APWC on CSE). IEEE; 2016. 34-39.

[51] Saha S, Ahmed KI, Mostafa R. Wavelet coherence based channel selection for classifying single trial motor imagery. In 2016 9th International Conference on Electrical and Computer Engineering (ICECE). IEEE; 2016. 467-470.

[52] Lu H, Eng HL, Guan C, Plataniotis KN, Venetsanopoulos AN. Regularized common spatial pattern with aggregation for EEG classification in small-sample setting. IEEE Transactions on Biomedical Engineering 2010; 57(12): 2936-2946.

[53] Narin A, Isler Y, Ozer M. Investigating the performance improvement of HRV indices in CHF using feature selection methods based on backward elimination and statistical significance. Computers in Biology and Medicine 2014; 45: 72-79.

[54] Isler Y. A detailed analysis of the effects of various combinations of heart rate variability indices in congestive heart failure (doctoral thesis). Izmir: Dokuz Eylul University; 2009.

[55] Mousa MA, El-Khoribi RA, Shoman ME. A novel brain computer interface based on principle component analysis. Procedia Computer Science 2016; 82: 49-56.

[56] Arteaga MV, Castiblanco JC, Mondragon IF, Colorado JD, Alvarado-Rojas, C. EMG-driven hand model based on the classification of individual finger movements. Biomedical Signal Processing and Control 2020; 58: 101834.

[57] Lee HS, Schreiner L, Jo SH, Sieghartsleitner S, Jordan M, Pretl H, *et al*. Individual finger movement decoding using a novel ultra-high-density electroencephalography-based brain-computer interface system. Frontiers in Neuroscience 2022; 16: 1009878.

[58] Mawase F, Cherry-Allen K, Xu J, Anaya M, Uehara S, Celnik P. Pushing the rehabilitation boundaries: Hand motor impairment can be reduced in chronic stroke. Neurorehabilitation and Neural Repair 2020; 34(8): 733-745.

[59] Bullock IM, Ma RR, Dollar AM. A hand-centric classification of human and robot dexterous manipulation. IEEE Transactions on Haptics 2013; 6(2) 129-144.

[60] Feix T, Romero J, Schmiedmayer HB, Dollar AM, Kragic D. The grasp taxonomy of human grasp types. IEEE Transactions on Human-Machine Systems 2016; 46(1): 66-77.

[61] Cutkosky MR. On grasp choice, grasp models, and the design of hands for manufacturing tasks. IEEE Transactions on Robotics and Automation 1989; 5(3): 269-279.

[62] Thrift AG, Thayabranathan T, Howard G, Howard VJ, Rothwell PM, Feigin VL, *et al*. Global stroke statistics. International Journal of Stroke 2017; 12(1): 13-32.

[63] Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, *et al*. Heart disease and stroke statistics-2018 update: a report from the American Heart Association. Circulation 2018; 137(12): e67-e492.

[64] Alazrai R, Alwanni H, Daoud MI. EEG-based BCI system for decoding finger movements within the same hand. Neuroscience letters 2019; 698: 113-120.

[65] Zahra HN, Zakaria H, Hermanto BR. Exploration of Pattern Recognition Methods for Motor Imagery EEG Signal with Convolutional Neural Network Approach. In Journal of Physics: Conference Series 2022; 2312(1): 012064.
[66] Kaya M, Binli MK, Ozbay E, Yanar H, Mishchenko Y. A large electroencephalographic motor imagery dataset for electroencephalographic brain computer interfaces. Scientific Data 2018; 5(1): 1-16.
[67] Mwata-Velu TY, Avina-Cervantes JG, Ruiz-Pinales J, Garcia-Calve TA, Gonzalez-Barbosa JB, Hurtado-Ramos JJ, *et al*. Improving Motor Imagery EEG Classification Based on Channel Selection Using a Deep Learning Architecture. Mathematics 2022; 10(13): 2302.
[68] Yahya N, Musa H, Ong ZY, Elamvazuthi I. Classification of motor functions from electroencephalogram (EEG) signals based on an integrated method comprised of common spatial pattern and wavelet transform framework. Sensors 2019; 19(22): 4878.
[69] Azizah RN, Zakaria H, Hermanto BR. Channels Selection for Pattern Recognition of Five Fingers Motor Imagery Electroencephalography Signals. In Journal of Physics: Conference Series 2022; 2312(1), 012019.
[70] Mwata-Velu TY, Avina-Cervantes JG, Cruz-Duarte JM, Rostro-Gonzalez H, Ruiz-Pinales J. Imaginary Finger Movements Decoding Using Empirical Mode Decomposition and a Stacked BiLSTM Architecture. Mathematics 2021; 9(24): 3297.
[71] Anam K, Nuh M, Al-Jumaily A. Comparison of EEG pattern recognition of motor imagery for finger movement classification. In 2019 6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI). IEEE; 2019. 24-27.
[72] Anam K, Bukhori S, Hanggara FS, Pratama M. Subject-independent Classification on Brain-Computer Interface using Autonomous Deep Learning for finger movement recognition. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE; 2020. 447-450.
[73] Kato M, Kanoga S, Hoshino T, Fukami T. Motor imagery classification of finger motions using multiclass CSP. In 2020 42nd Annual International Conference

of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE; 2020. 2991-2994.

[74] Limbaga NJ, Mallari KL, Yeung NR, Monje JC. Development of an EEG-based Brain-Controlled System for a Virtual Prosthetic Hand. In 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE; 2022.1714-1717.

[75] Degirmenci M, Yuce YK, Isler Y. Classification of Finger Movements from Statistically-Significant Time-Domain EEG Features. Journal of the Faculty of Engineering and Architecture of Gazi University, Accepted, 2023.

[76] Brunner C, Leeb R, Muller-Putz G, Schlogl A, Pfurtscheller G. BCI Competition 2008–Graz data set A. Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology 2008; 16: 1-6. Available from: http://www.bbci.de/competition/iv/desc_2a.pdf.

[77] Geethanjali P, Mohan YK, Sen J. Time domain feature extraction and classification of EEG data for brain computer interface. In 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery. IEEE; 2012. 1136-1139.

[78] Boubchir L, Daachi B, Pangracious V. A review of feature extraction for EEG epileptic seizure detection and classification. 2017 40th International Conference on Telecommunications and Signal Processing (TSP). IEEE; 2017. 456-460.

[79] Sayilgan E. Evaluating Steady-State Visually-Evoked Potentials Using Ensemble Learning Methods (doctoral thesis). Izmir: Izmir Katip Celebi University; 2020.

[80] Khosla A, Khandnor P, Chand T. A comparative analysis of signal processing and classification methods for different applications based on EEG signals. Biocybernetics and Biomedical Engineering 2020; 40(2): 649-690.

[81] Degirmenci M, Yuce YK, Isler Y. Motor imaginary task classification using statistically significant time domain and frequency domain eeg features. Journal of Intelligent Systems with Applications 2022; 5(1): 49-54.

[82] Safi MS, Safi SMM. Early detection of Alzheimer's disease from EEG signals using hjorth parameters. Biomedical Signal Processing and Control 2021; 65: 102338.

[83] Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 1989; 11(7): 674-693.

[84] Akansu N, Tazebay MV, Medley MJ, Das PK. Wavelet and subband transforms: fundamentals and communication applications. IEEE Communications Magazine 1997; 35(12): 104-115.

[85] Abbate A, Decusatis CM, Das PK. Wavelets and subbands: fundamentals and application: Birkhéuser Boston; 1995.

[86] Gandhi T, Panigrahi KB, Anand S. A comparative study of wavelet families for EEG signal classification. Neurocomputing 2011; 74(17): 3051-3057.

[87] Alickovic E, Kevric J, Subasi A. Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction. Biomedical Signal Processing and Control 2018; 39: 94-102.

[88] Chen D, Wan S, Xiang J, Bao FS. A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG. PLoS ONE 2017; 12(3): e0173138.

[89] Gao Z, Li S, Cai Q, Dang W, Yang Y, Mu C, Hui P. Relative wavelet entropy complex network for improving EEG-based fatigue driving classification. IEEE Transactions on Instrumentation and Measurement 2019; 68(7): 2491-2497.

[90] Amin HU, Malik AS, Ahmad RF, Badruddinb N, Kamel N, Hussain M, Chooi WT. Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. Australasian Physical & Engineering Sciences in Medicine 2015; 38(2015): 139-149.

[91] Al-Fahoum AS, Al-Fraihat AA. Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. International Scholarly Research Notices (ISRN) Neuroscience 2014; 7(730218).

[92] Heidari H, Einalou Z. SSVEP extraction applying wavelet transform and decision tree with Bayes classification. International Clinical Neuroscience Journal (ICNSJ) 2017; 4(3): 91-97.

[93] Frei MG, Osorio I. Intrinsic time-scale decomposition: time–frequency–energy analysis and real-time filtering of non-stationary signals. Proceedings of the

Royal Society A: Mathematical, Physical and Engineering Sciences 2007; 463(2078): 321-342.

[94] Voznesensky A, Kaplun D. Adaptive signal processing algorithms based on EMD and ITD. IEEE Access 2019; 7: 171313-171321.

[95] Martis RJ, Acharya UR, Tan JH, Petznick A, Tong L, Chua CK, *et al*. Application of intrinsic time-scale decomposition (ITD) to EEG signals for automated seizure prediction. International Journal of Neural Systems 2013; 23(05): 1350023.

[96] Iscan Z, Dokur Z, Demiralp T. Classification of electroencephalogram signals with combined time and frequency features. Expert Systems with Applications 2011; 38(8): 10499-10505.

[97] Cura OK, Atli SK, Akan A, Attention deficit hyperactivity disorder recognition based on intrinsic time-scale decomposition of EEG signals. Biomedical Signal Processing and Control 2023; 81: 104512.

[98] Degirmenci M, Ozdemir MA, Sadighzadeh R, Akan A. Emotion recognition from EEG signals by using empirical mode decomposition. In 2018 Medical Technologies National Congress (TIPTEKNO). IEEE: 2018. 1-4.

[99] Hjorth B. EEG analysis based on time domain properties. Electroencephalography and Clinical Neurophysiology 1970; 29(3): 306-310.

[100] Yu M, Fang M. Feature extraction of rolling bearing multiple faults based on correlation coefficient and Hjorth parameter. ISA transactions 2022; 129: 442-458.

[101] Higuchi T. Approach to an irregular time series on the basis of the fractal theory. Physica D: Nonlinear Phenomena 1988; 31(2): 277-283.

[102] Isler Y, Kuntalp M. Diagnosis of congestive heart failure patients using Poincare measures derived from ECG signals. In 2009 14th National Biomedical Engineering Meeting (BIYOMUT). IEEE; 2009. 1-4.

[103] Isler Y, Narin A, Ozer M, Perc M. Multi-stage classification of congestive heart failure based on short-term heart rate variability. Chaos, Solitons & Fractals 2019; 118: 145-151.

[104] Brennan M, Palaniswam M, Kamen P. Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability?. IEEE Transactions on Biomedical Engineering 2001; 48(11): 1342-1347.

[105] Smith AL, Reynolds KJ, Owen H. Correlated Poincare indices for measuring heart rate variability. Australasian Physical & Engineering Sciences in Medicine 2007; 30(4): 336-341.

[106] Stein PK, Le Q, Domitrovich PP. Development of more erratic heart rate patterns is associated with mortality post-MI. Journal of Electrocardiology 2008; 41(2): 110-115.

[107] Duda RO, Hart PE, Stork DG. Pattern Classification: John Wiley & Sons; 2001.

[108] Liuand H, Motoda H. Feature selection for knowledge discovery and data mining: Kluwer Academic Publishers; 1998.

[109] Degirmenci M, Yuce YK, Perc M, Isler Y. Statistically significant features improve binary and multiple motor imagery task predictions from EEGs. Frontiers in Human Neuroscience 2023; 17: 1223307.

[110] Sahin A, Ozcan N, Nur G. Ovarian cancer prediction using PCA, K-PCA, ICA and random forest. Jornal of Intelligent Systems with Applications 2021; 4(2): 103-108.

[111] Yesilkaya B, Perc M, Isler Y. Manifold learning methods for the diagnosis of ovarian cancer. Journal of Computational Science 2022; 63: 101775.

[112] Tzallas AT, Tsipouras MG, Fotiadis DI. Epileptic seizure detection in EEGs using time–frequency analysis. IEEE Transactions on Information Technology in Biomedicine 2009; 13(5): 703-710.

[113] Lotte F. Study of electroencephalographic signal processing and classification techniques towards the use of brain-computer interfaces in virtual reality applications. Human-Computer Interaction [cs.HC]. INSA de Rennes, 2008.

[114] Alpaydin E. Introduction to Machine Learning: MIT Press; 2004.

[115] James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning: Springer; 2013.

[116] Samiee K, Kiranyaz S, Gabbouj M, Saramaki T. Long-term epileptic EEG classification via 2D mapping and textural features. Expert Systems with Applications 2015; 42(20): 7175-7185.

[117] James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning: Springer; 2013.

[118] Rokach L. Ensemble-based classifiers. Artificial Intelligence Review. 2010; 33: 1-39.

[119] Polikar R. Ensemble learning. Scholarpedia. 2009; 4(1): 2776.

[120] Necati Demir. Ensemble methods: Elegant techniques to produce improved machine learning results. Available from: https://www.toptal.com/machinelearning/ensemble-methods-machine-learning. [Accessed 3rd May 2019].

[121] Lei D, Tang J, Li Z, Wu Y. Using low-rank approximations to speed up kernel logistic regression algorithm. IEEE Access 2019; 7: 84242-84252.

[122] Maji S, Berg AC, Malik J. Classification using intersection kernel support vector machines is efficient. In 2008 IEEE Conference on Computer Vision and Pattern Recognition. IEEE: 2008. 1-8.

[123] Ozdemir MA, Degirmenci M, Izci E, Akan A. EEG-based emotion recognition with deep convolutional neural networks. Biomedical Engineering/Biomedizinische Technik 2021; 66(1): 43-57.

[124] Degirmenci M, Ozdemir MA, Izci E, Akan A, Arrhythmic heartbeat classification using 2D convolutional neural networks. Irbm 2022; 43(5): 422-433.

[125] Narin A, Isler Y. Detection of new coronavirus disease from chest x-ray images using pre-trained convolutional neural networks. Journal of the Faculty of Engineering and Architecture of Gazi University 2021; 36(4): 2095-2107.

[126] Richard MD, Lippmann RP. Neural network classifiers estimate Bayesian a posteriori probabilities. Neural computation 1991; 3(4), 461-483.

[127] Pan S, Iplikci S, Warwick K, Aziz TZ. Parkinson's Disease tremor classification–A comparison between Support Vector Machines and neural networks. Expert Systems with Applications 2012; 39(12): 10764-10771.

[128] Xu G, Shen X, Chen S, Zong Y, Zhang C, Yue H, *et al*. A deep transfer convolutional neural network framework for EEG signal classification. IEEE Access 2019; 7: 112767-112776.

[129] Zhao D, Tang F, Si B, Feng X. Learning joint space–time–frequency features for EEG decoding on small labeled data. Neural Networks 2019; 114: 67-77.

[130] Sakhavi S, Guan C, Yan S. Parallel convolutional-linear neural network for motor imagery classification. In 2015 23rd European Signal Processing Conference (EUSIPCO). IEEE: 2015. 2736-2740.

[131] Jusas V, Samuvel SG. Classification of motor imagery using combination of feature extraction and reduction methods for brain-computer interface. Information Technology and Control 2019; 48(2): 225-234.

[132] Nguyen T, Hettiarachchi I, Khatami A, Gordon-Brown, Lim CP, Nahavandi, S. Classification of multi-class BCI data by common spatial pattern and fuzzy system. IEEE Access 2018; 6: 27873-27884.

[133] Lindig-Leon C, Bougrain L. A multi-label classification method for detection of combined motor imageries. In 2015 IEEE International Conference on Systems, Man, and Cybernetics. IEEE: 2015. 3128-3133.

[134] Lee SB, Kim HJ, Kim H, Jeong JH, Lee SW, Kim DJ. Comparative analysis of features extracted from EEG spatial, spectral and temporal domains for binary and multiclass motor imagery classification. Information Sciences 2019; 502: 190-200.

Appendices

Appendix A

Publications from the Thesis

Conference Papers

1. Degirmenci M, Yuce YK, Isler Y (2022). Motor Imginary Task Classification using Statistically Significant Time-Domain EEG Features. 2022 30th Signal Processing and Communications Applications Conference (SIU), Doi: 10.1109/SIU55565.2022.9864745 (Tam Metin Bildiri/Sözlü Sunum)

2. Degirmenci M, Yuce YK, Isler Y (2022). Classification of Multi-Class Motor Imaginary Tasks using Poincare Measurements Extracted from EEG Signals. 2022 5th International Conference on Medical Devices (ICMD'2022), (Özet Bildiri/Sözlü Sunum)

3. Degirmenci M, Yuce YK, Isler Y (2022). Motor Imaginary Task Classification using Statistically Significant Time Domain and Frequency Domain EEG Features. 2022 5th International Conference on Medical Devices (ICMD'2022), (Özet Bildiri/Sözlü Sunum)

4. Degirmenci M, Sayilgan E, Yuce YK, Isler Y. Evaluation of Wigner-Ville Distribution Features to Estimate Steady-State Visual Evoked Potentials' Stimulation Frequency. 2021 4th International Conference on Medical Devices (ICMD'2021), 27 (Özet Bildiri/Sözlü Sunum)

Journal Articles

1. Degirmenci M, Yuce YK, Perc M, Isler Y. Statistically Significant Features Improve Binary and Multiple Motor Imagery Tasks Predictions from EEGs. Frontiers in Human Neuroscience 2023; 17: 1223307. Doi: https://doi.org/10.3389/fnhum.2023.1223307.

2. Degirmenci M, Yuce YK, Isler Y. Classification of Finger Movements from Statistically Significant Time-Domain EEG Features. Journal of the Faculty of Engineering and Architecture of Gazi University, ACCEPTED, 2023.

3. Degirmenci M, Yuce YK, Isler Y. Classification of Multi-Class Motor Imaginary Tasks using Poincare Measurements Extracted from EEG Signals. Journal of Intelligent Systems with Applications 2022; 5(2): 74-78. Doi: 10.54856/jiswa.202212204.

4. Degirmenci M, Yuce YK, Isler Y. Motor Imaginary Task Classification using Statistically Significant Time Domain and Frequency Domain EEG Features. Journal of Intelligent Systems with Applications 2022; 5(1): 49-54. Doi: 10.54856/jiswa.202205203.

Projects

1. Investigation of the Effects of Various Time-Frequency Representations on EEG-Based Motor Imaginary Task Classification. Supported by Scientific Projects Council of Izmir Katip Celebi University (2023-TDR-FEBE-0002).

Curriculum Vitae

Name Surname : Murside Degirmenci

Education:

2019–2024 İzmir Kâtip Çelebi University,
PhD in Biomedical Technologies

2013–2018 İzmir Kâtip Çelebi University,
B.Sc in Biomedical Engineering

Journal Articles:

1. Degirmenci M, Yuce YK, Perc M, Isler Y. Statistically Significant Features Improve Binary and Multiple Motor Imagery Tasks Predictions from EEGs. Frontiers in Human Neuroscience 2023; 17: 1223307. Doi: https://doi.org/10.3389/fnhum.2023.1223307

2. Degirmenci M, Yuce YK, Isler Y. Classification of Finger Movements from Statistically Significant Time-Domain EEG Features. Journal of the Faculty of Engineering and Architecture of Gazi University, ACCEPTED, 2023.

3. Degirmenci M, Ozdemir MA, Izci E, Akan A. Arrhythmic Heartbeat Classification using 2D Convolutional Neural Networks. Innovation and Research in Biomedical Engineering (IRBM) 2022; 43(5): 422-433. Doi: https://doi.org/10.1016/j.irbm.2021.04.002

4. Ozdemir M, **Degirmenci M,** Izci E, Akan A. EEG-Based Emotion Recognition with Deep Convolutional Neural Networks. Biomedical Engineering/Biomedizinische Technik 2021; 66(1): 43-57. Doi: https://doi.org/10.1515/bmt-2019-0306

5. Degirmenci M, Yuce YK, Isler Y. Classification of Multi-Class Motor Imaginary Tasks using Poincare Measurements Extracted from EEG Signals. Journal of Intelligent Systems with Applications 2022; 5(2): 74-78. Doi: 10.54856/jiswa.202212204

6. Degirmenci M, Yuce YK, Isler Y. Motor Imaginary Task Classification using Statistically Significant Time Domain and Frequency Domain EEG Features. Journal of Intelligent Systems with Applications 2022; 5(1): 49-54. Doi: 10.54856/jiswa.202205203

7. Degirmenci M, Sayilgan E, Yuce YK, Isler Y. Evaluation of Wigner-Ville Distribution Features to Estimate Steady-State Visual Evoked Potentials' Stimulation Frequency. Journal of Intelligent Systems with Applications 2021; 4(2): 133-136. Doi: 10.54856/jiswa.202112178

Conference Papers:

1. Izci E, Ozdemir MA, **Degirmenci M,** Akan A (2019). Cardiac Arrhythmia Detection from 2d ECG Images by using Deep Learning Technique. 2019 Medical Technologies Congress (TIPTEKNO), Doi: 10.1109/TIPTEKNO.2019.8895011 (Tam Metin Bildiri/Sözlü Sunum)

2. Izci E, **Degirmenci M,** Ozdemir MA, Akan A (2020). ECG Arrhythmia Detection with Deep Learning. 2020 28th Signal Processing and Communications Applications Conference (SIU), Doi: 10.1109/SIU49456.2020.9302219 (Tam Metin Bildiri/Sözlü Sunum)

3. Ozdemir MA, **Degirmenci M,** Guren O, Akan A (2019). EEG based Emotional State Estimation using 2-D Deep Learning Technique. 2019 Medical Technologies Congress (TIPTEKNO), Doi: 10.1109/TIPTEKNO.2019.8895158 (Tam Metin Bildiri/Sözlü Sunum)

4. Degirmenci M, Ozdemir MA, Sadighzadeh R, Akan A (2018). Emotion Recognition from EEG Signals by using Empirical Mode Decomposition. 2018

Medical Technologies Congress (TIPTEKNO), Doi: 10.1109/TIPTEKNO.2018.8597061 (Tam Metin Bildiri/Sözlü Sunum)

5. Degirmenci M, Akan A (2020). EEG based Epileptic Seizures Detection using Intrinsic Time-Scale Decomposition. 2020 Medical Technologies Congress (TIPTEKNO), Doi: 10.1109/TIPTEKNO50054.2020.9299262 (Tam Metin Bildiri/Sözlü Sunum)

6. Degirmenci M, Yuce YK, Isler Y (2022). Motor Imginary Task Classification using Statistically Significant Time-Domain EEG Features. 2022 30th Signal Processing and Communications Applications Conference (SIU), Doi: 10.1109/SIU55565.2022.9864745 (Tam Metin Bildiri/Sözlü Sunum)

7. Degirmenci M, Sayilgan E, Isler Y (2021). Extracting the Steady-State Visual Evoked-Potentials Using Wigner-Ville Distribution. 2021 3rd International Conference of Applied Sciences, Engineering and Mathematics (IBU-ICASEM 2021), 20 (Özet Bildiri/Sözlü Sunum)

8. Degirmenci M, Yuce YK, Isler Y (2022). Classification of Multi-Class Motor Imaginary Tasks using Poincare Measurements Extracted from EEG Signals. 2022 5th International Conference on Medical Devices (ICMD'2022), (Özet Bildiri/Sözlü Sunum)

9. Degirmenci M, Yuce YK, Isler Y (2022). Motor Imaginary Task Classification using Statistically Significant Time Domain and Frequency Domain EEG Features. 2022 5th International Conference on Medical Devices (ICMD'2022), (Özet Bildiri/Sözlü Sunum)

10. Degirmenci M, Sayilgan E, Yuce YK, Isler Y. Evaluation of Wigner-Ville Distribution Features to Estimate Steady-State Visual Evoked Potentials’ Stimulation Frequency. 2021 4th International Conference on Medical Devices (ICMD'2021), 27 (Özet Bildiri/Sözlü Sunum)