BEYOND HUMAN: A MACHINE LEARNING APPROACH TO EEG-BASED EMOTION CLASSIFICATION

MULTI-CLASS CLASSIFICATION EMPLOYING KNN AND SVM

ESTHER I. SENNEKER

THESIS SUBMITTED IN FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER IN NEW MEDIA DESIGN

DEPARTMENT OF
COMMUNICATION & INFORMATION SCIENCES
SCHOOL OF HUMANITIES AND DIGITAL SCIENCES
TILBURG UNIVERSITY

STUDENT NUMBER

2038724

COMMITTEE

Dr. David Peeters
Dr. Tessa van Leeuwen

LOCATION

Tilburg University
School of Humanities and Digital Sciences
Department of Communication &
Information Sciences
Tilburg, The Netherlands

DATE

June 20, 2024

WORD COUNT

9875 words

ACKNOWLEDGMENTS

I want to express my sincerest gratitude to Dr. David Peeters for his help during the process of creating this master thesis. He provided me with his insights and feedback, which greatly contributed to the quality of this paper. I felt greatly supported and encouraged from the beginning till the completion, and I want to thank him for that.

TECHNOLOGY STATEMENT

All text in this thesis is written by the author named on the title page. No content was copied from other sources without proper citation. ChatGPT was used to provide the author with word synonyms and feedback on text created by the author. A spelling and grammar checker by the name of LanguageTool was used during the creation of this thesis. An online tool named Overleaf was used to format this thesis. ChatGPT and AskCodi were used to aid the researcher with coding problems during the data processing of this thesis. No other generative tools or services were used to create this thesis.

ESTHER I. SENNEKER

Abstract

This thesis contributes to the evolving field of EEG-based emotion classification by investigating the efficacy of Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) models. The performance of the SVM and KNN was compared, with accuracies of respectively 87% and 85% in classifying 15 distinct emotions across extended durations. Subsequent analyses explored the impact of unique data points from multiple participants on model performance, and resulted in no significant difference in performance. The influence of electrode exclusion on classification accuracy was investigated, and the difference in accuracy of the models was not significant. Findings suggest that SVM exhibits superior classification capabilities, particularly in complex, high-dimensional EEG datasets. Moreover, the study underscores the importance of methodological considerations, such as feature selection and model optimization, in enhancing emotion classification accuracy.

KEYWORDS: EEG; Emotion classification; Machine learning; KNN; SVM; Physiological signals.

1 INTRODUCTION

Human emotion has been widely researched for centuries now as it plays a fundamental role in human decision-making processes, shaping experiences, interactions, and human motivation (Andersen & Guerrero, 1996; Darwin, 1859). Previous studies have developed and evaluated numerous psychological models concerning the causes and effects of emotions (Scherer et al., 2000). Many of the fundamental differences between these models originate from the definition of emotion they are based upon. This paper makes use of the definition provided by Scherer et al. (2000, p. 140) for emotion, which states that emotion is:

> *"A relatively brief episode of synchronized responses by all or most organismic subsystems to the evaluation of an external or internal event as being of major significance (e.g., anger, sadness, joy, fear, shame, pride, elation, desperation)"*

Understanding and accurately interpreting emotions is essential for effective communication, empathy, and social cohesion within human interaction (Frijda, 2007). Furthermore, emotions can be the consequence of or significantly impact various aspects of human life, including mental health, physical well-being, and the overall quality of life (Frijda, 2007). Thus, understanding emotion is important for individuals as well as for collective groups of people interacting with each other. Researching emotion enhances our understanding of the human psyche and holds the potential to improve mental health interventions, enhance human-human interactions, enhance human-computer interactions, and positively affect the quality of human life (Kuppens & Verduyn, 2017). However, measuring and understanding emotions has proven to be difficult. Their subjective nature and the human ability to mask them (social desirability bias), leaves room for different interpretations (Niedenthal & Ric, 2017). Many research methods have been employed over the years to try to overcome these difficulties, with a recent increase in the use of computer algorithms to better detect, recognize, and classify human emotion (Saxena, Khanna, & Gupta, 2020).

Emotion detection and classification through the use of Artificial Intelligence have become popular topics of research within the scientific field (Saxena et al., 2020). Emotion detection primarily relies on four main approaches: distinguishing facial expressions, analyzing variations in speech signals, interpreting physiological signals, and analysis of textual semantics (Saxena et al., 2020). This study will focus on interpreting physiological signals. A non-invasive method used to record physiological signals is electroencephalography (EEG). It records the electrical activity in a subject's brain through electrodes placed on their head, which allows for it to be analyzed through the use of machine learning (Siuly, Li, & Zhang, 2016). The type of machine learning analysis that this paper concerns itself with is classification. Classification analysis is a method used to categorize data into different classes or groups based on certain features or attributes (Wang & Wang, 2021). The aim is to develop a model that can accurately assign new data points to their appropriate class based on patterns learned from its training dataset (Wang & Wang, 2021).

The current study makes use of Support Vector Machines (SVM) and K-Nearest Neighbor (KNN) models, since they yielded amongst the highest accuracies, with a 93% and 92% respectively when trained and tested on the Dataset for Emotion Analysis by Wang and Wang (2021) and Yu and Wang (2022). KNN is a machine learning algorithm that classifies data points based on the majority class of their nearest neighbors, making predictions by finding the most similar instances in the training data. SVM is a supervised learning algorithm that finds the optimal hyperplane separating different classes in the feature space, maximizing the margin between classes, to classify new data points (Wang & Wang, 2021; Yu & Wang, 2022).

The previously mentioned high accuracies for the KNN and SMV models, imply that these two machine learning models are suitable for EEG based classification. However, it is important to note that these accuracies are specifically based on the models' performance on the DEAP. This entails that the characteristics and ecological validity of the experiment that provided the data in the DEAP are of importance when considering the generalizability of any models trained on it. Many studies that have been conducted in the past on emotion classification use the publicly available DEAP, SEED, and DREAMER (Cimtay & Ekmekcioglu, 2020; Cui et al., 2020; Placidi, Di Giamberardino, Petracca, Spezialetti, & Iacoviello, 2016). This paper does not make use of those datasets due to an abundance of past research already providing the emotion classification results, and due to their drawbacks. It is important to understand the drawbacks of using these three datasets, which will be discussed in section 2.3.

Even though extensive research has been done on the topic of emotion classification, a highly accurate generalizable model has not been developed as of today. The aim of this paper is to aid in the development of such a model as it could be employed for various purposes that play into human emotion. Examples of such purposes are the development of more capable social robots, and new media applications that optimize user experience. Being able to perform user product testing, while not relying on self-reported emotions, can remove ambiguity originating from social desirability bias and emotional recall bias (Levine, Lench, & Safer, 2009). To help achieve this goal, the current study bases its classification models on a recently released EEG dataset (Onton & Makeig, 2022). This dataset contains rich EEG data on 15 different emotions that were felt across multiple minutes, and its experimental setup will be compared to the previously datasets in section 2.3.

In addition to emotion classification based on that dataset, this study concerns itself with the impact of test-train data splitting on accuracy. Research has shown that patterns of brain activity associated with different emotions are more consistent within individuals than between different individuals experiencing the same emotion (Hsu, Lin, Onton, Jung, & Makeig, 2022). For example, the neural activity of participant 1 while feeling sad can more closely resemble the neural activity of participant 1 feeling happy than the neural activity of participant 2 feeling sad. When investigated, the results can provide insights into the generalizability and robustness of the models, ultimately enhancing the validity and reliability of the emotion classification results and providing future researchers with additional information on the effect of different train and tests splits.

Lastly, the current research tests a novel subset of electrodes and compares the accuracy to the accuracy of the models when trained on the whole dataset. Pessoa (2017) suggests that only a subset of brain structures are involved with emotion. Based on this knowledge, research can be conducted about excluding EEG data from electrodes not detecting signals originating from the mentioned brain structures from the classification process. If accuracy turns out to be unaffected by the exclusion of electrodes, a generalizable emotion classification model could be less costly, both computationally and financially. The following research questions were formulated to accomplish the aims of this study:

RQ: ***"How accurately can Support Vector Machines and K-Nearest Neighbor models classify 15 different emotions felt across multiple minutes from electroencephalographic data?"***

To answer the additional questions mentioned above, the following sub-questions were formulated:

SQ1 *"Which of the two classification models yields a higher accuracy?"*

SQ2 *"Is the emotion classification accuracy of Support Vector Machines and K-Nearest Neighbor models affected when unique data from the same participant is included in the test and train dataset?"*

SQ3 *"Is the emotion classification accuracy of Support Vector Machines and K-Nearest Neighbor models affected by the exclusion of electrodes not detecting signals originating from brain structures involved with emotion?"*

Novel to this research is the use of a dataset with improved emotion labels that more accurately represent the emotional state of the participant at any given time during the experiment. This is a modern dataset with improved quality due to its use of updated equipment, the vast amount of data points, and the implementation of instant emotion labelling instead of relying on post-task self-reporting of emotions. Furthermore, the inclusion of research into effective test-train data splitting and subset classification can aid future researchers with allocating their resources when utilizing classification. As mentioned before, emotions can be affected and affect many aspects of human life. A better understanding of them may lead to improved mental health interventions, social interactions, socially improved artificially intelligent assistance, and robotic aids for mental and physical healthcare. In sum, this study's results can aid future researchers with allocating their computing power and in the developemtn of improved robot assistance in mental healthcare.

2 THEORETICAL FRAMEWORK

2.1 *Researching Emotion*

Human emotion has been a subject of study a long time, with one of the earliest theories of emotion leading back to the late 19th century (Dewey, 1895). This theory by William James and Carl Lange suggests that emotions arise from physiological responses to stimuli in the environment, and that one's emotional reaction is dependent on their interpretation of those physiological responses (Dewey, 1895). Many revisions of this theory and new theories have been proposed since then, however, these still express the subjectivity of experiencing emotions (Plutchik & Kellerman, 2013). Emotional responses might be subjective, however, the physiological reactions involved in the process of feeling emotion can be measured objectively. Modern equipment has allowed researchers to record heart rates, blood pressure, brain activity, eye movement, etc. Recorded physiological signals are less susceptible to emotional recall bias, social desirability bias, subjective interpretation, and cultural biases (Levine et al., 2009). The mentioned factors can interfere with the quality and generalizability of results, as they can happen consciously and unconsciously. Physiological responses are valuable data on their own. training machine learning techniques on physiological data allows for the analysis of millions of these data points in one model Zhang and Chen (2020). When enough labeled data is presented to a classification algorithm, it can discover underlying patterns in the data Zhang and Chen (2020). These underlying patterns enable models to classify new unlabeled data points to one of the classes. This method of research has increased in popularity as it can provide researchers with more information on human experiences such as emotion Zhang and Chen (2020).

2.2 *Classification Models*

When considering EEG classification, it is essential to investigate and consider which classification models to use. Wang and Wang (2021) reviewed machine learning models that can be used for emotion classification with EEG data. The DEAP dataset was used in their paper to train and test all classification models. Wang and Wang (2021) states that emotion classification based on electroencephalographic data has mostly been performed using the following models: Support Vector Machine (SVM), Naive Bayes (NB), and k-Nearest Neighbor (KNN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long Short-term Memory Neural Network (LSTM). The paper presented the accuracy results for all previously mentioned models, which ranged between 50-97% (Wang & Wang, 2021; Yu & Wang, 2022). The KNN and SVM models yielded the highest accuracies of 92% and 93% respectively.

Additionally, Amin, Mumtaz, Subhani, Saad, and Malik (2017) tested the performance of different machine learning algorithms in classifying EEG signals recorded during various cognitive tasks. Specifically, the study explored the effectiveness of pattern recognition techniques and feature extraction methods in discriminating EEG signals associated with complex cognitive tasks from those recorded during baseline tasks. Complex cognitive tasks included tasks such as mentally composing a letter to a loved one without vocalizing any part, while an example baseline task was having the eyes open. In the paper, EEG signals were broken down into different frequency components, which they refer to as coefficients. The results of this study support the notion that KNN and SVM are suitable models to use for EEG based classification on large datasets. Notably, KNN yielded an accuracy of 98.39% for detailed coefficients, while SVM achieved an accuracy of 99.11% for low-frequency coefficients. Furthermore, detailed coefficients derived from the sub-band range (3.90–7.81 Hz) yielded accuracies of 98.57% for SVM and 93.33% for KNN, further confirming the effectiveness of these models in discerning subtle cognitive patterns embedded within EEG data. These findings validate the suitability of KNN and SVM for EEG-based classification tasks and highlight their potential utility in real-world applications requiring precise and reliable cognitive task classification methodologies.

While significant strides have been made in emotion classification using machine learning models like KNN and SVM, there are still notable gaps in the scientific literature that warrant further investigation. As presented by Wang and Wang (2021) and Yu and Wang (2022), classification accuracies on the DEAP, SEED, and DREAMER datasets are high. However, these datasets have drawbacks which warrant additional research on EEG-based emotion classification with a dataset that addresses these shortcomings. These specific drawbacks will be discussed in the next section. In sum, the findings from Amin et al. (2017) and Wang and Wang (2021), suggest that the new dataset should be trained and tested with the KNN and SVM models to achieve maximum accuracies.

2.3 *EEG-based Emotion Classification*

To aid in the development of a highly accurate generalized emotion classification model, reflection on relevant past scientific literature is needed. Past research on emotion classification has often made use of EEG (Alarcao & Fonseca, 2017). There are several advantages of using EEG that led to its popularity. Firstly, EEG provides excellent temporal resolution, enabling researchers to capture the rapid dynamics of brain activity associated with emotional processing (Alarcao & Fonseca, 2017). This temporal precision allows for the examination of moment-to-moment changes in emotional states, providing insights into the temporal unfolding of emotions. Secondly, EEG is non-invasive, making it relatively safe and comfortable for participants, which is essential for studying emotional responses that may be influenced by stress or discomfort (Alarcao & Fonseca, 2017). Thirdly, EEG is relatively cost-effective and portable compared to other neuroimaging techniques, allowing for large-scale studies and longitudinal assessments of emotional functioning (Alarcao & Fonseca, 2017). Lastly, EEG is well-suited for capturing the oscillatory patterns and event-related potentials associated with emotional processing, providing rich data for analysis (Alarcao & Fonseca, 2017). Overall, these advantages have contributed to EEG's widespread use in emotion classification research, enabling researchers to explore the neural correlates of emotions with high precision and sensitivity.

Many past scientific studies regarding emotion classification made use of publicly available datasets. The most popular datasets are called DEAP, SEED, and DREAMER. An understanding of these datasets is needed to fully assess the implications of past classification accuracies based on their data. Firstly, the DEAP's experiment, conducted in 2012, consists of the data from 32 participants who were tasked to watch 40 one-minute clips from music videos (Koelstra et al., 2012). The participants subsequently filled out the Russell's valence-arousal scale, on which they plotted their emotion using a two-dimensional coordinate system. Valence is represented on the vertical axis (ranging from positive to negative) and arousal represented on the horizontal axis (ranging from low to high). Emotions can be plotted within this space based on their respective valence and arousal levels, with sadness being depicted as low-valence and low-arousal and happiness ad high-valence, high-arousal (Koelstra et al., 2012). The DEAP was recorded 12 years ago and, while it is valuable data, there are modern datasets containing more extensive and higher-quality EEG recordings (Hsu et al., 2022). The DEAP dataset recorded the data from 32 electrodes, while modern equipment allows for up to 256 electrodes (Siuly et al., 2016). The ecological validity of this experiment can also be questioned, since all emotions were induced by watching short music videos in a lab setting. The limited variety and complexity of these stimuli may not fully capture the breadth of emotional experiences encountered in real-life scenarios. As reported by Holm, Kaakinen, Forsström, and Surakka (2021), experienced emotional arousal and valence of watching something can differ from experiencing it first-hand, or with some involvement of the participant. Additionally, the induced emotions were labeled after the emotions had already subsided, allowing for emotional recall bias to affect the quality of labels.

The SEED dataset, documented by Duan, Zhu, and Lu (2013) and Zheng and Lu (2017), is an EEG dataset assembled by the Brain-like Computing and Machine Intelligence laboratory (BCMI). It consists of EEG recordings collected from 15 healthy subjects while they performed five cognitive tasks: resting state, auditory oddball, eye-closed visual stimulus, eye-open visual stimulus, and the motor imagery task (Duan et al., 2013). The EEG data was recorded using a 62-channel EEG system at a sampling rate of 500 Hz. Additionally, each subject completed a self-assessment manikin (SAM) questionnaire after each session to report their emotional states regarding the arousal (intensity), valence (positive or negative nature), and dominance (feeling of personal control in the emotional state)(Duan et al., 2013). However, like the DEAP, this dataset relies on participants' subjective ratings on the valence-arousal scale after an emotion has already subsided, introducing potential biases such as emotional recall bias (Levine et al., 2009). Additionally, this dataset suffers from a limited number of participants, which can limit the generalizability of findings.

Lastly, the DREAMER dataset, presented by Katsigiannis and Ramzan (2017), is comprised of EEG signals, along with self-reported emotional annotations, collected from 23 participants while they watched a series of 18 movie clips designed to induce various emotions. Each participant provided continuous ratings of arousal and valence throughout the movie viewing session, allowing for the alignment of emotional responses with specific moments in the videos and eliminating the factor of emotional recall bias. Additionally, participants completed post-viewing questionnaires to provide further insights into their emotional experiences. Limitations of the DREAMER are similar to the DEAP and SEED. It suffers from a lack of variety and complexity of emotional stimuli, which can impact the ecological validity of the dataset. Furthermore, it has a limited number of participants and relies on the participants' subjective rating of their emotion to label data.

While these datasets have been invaluable for advancing research in emotion recognition, efforts to address their shortcomings and enhance their methods of inducing emotions are warranted when aiming for more robust and generalizable findings. There are recent studies that make use of new datasets. A paper by Hsu et al. (2022) uses a dataset from 2022, that contains the EEG recordings of 34 participants in a self-paced emotion imagination task. During this experiment, participants were induced into a sequence of 15 emotional states (e.g. jealousy, love, and fear) by listening to recorded voice narratives with their eyes closed and engaging in personal imagination. Once the target emotion was felt by the participant they were tasked to press a handheld button, this could communicate it to the researchers in real time while causing minimal noise on the EEG recordings. The use of this button meant that participants did not have to self-report on their emotions afterward, but could instead indicate it throughout the experiment without jeopardizing data quality. The experiment resulted in a dataset consisting of 3-5 minute recordings for each emotion per participant. The EEG equipment recorded data through 256 electrodes, which provides classification models with rich data for all 15 emotions (Hsu et al., 2022). The results and implications of the study by Hsu et al. (2022) will be discussed in section 2.6 of this paper. The current research makes use of the dataset provided by Hsu et al. (2022) to train and test its classification model.

2.4 *Classification with KKN*

The KNN model is a simple yet highly effective machine learning algorithm used extensively for classification tasks (Kataria & Singh, 2013). KNN does the following when classifying a new data point: it calculates the distance between this point and all other data points in the training set. These distances are typically computed using metrics such as Euclidean distance. Once distances are determined, the algorithm identifies the majority class among the point's K closest neighbors, where K is a user-defined parameter, in the feature space, and identifies it to the new data point (Kataria & Singh, 2013).

This proximity-based approach makes KNN particularly adept at capturing intricate patterns present in complex datasets, such as those derived from EEG recordings (Wang & Wang, 2021). When considering past classification studies, the value of K commonly used ranges from 3 to 10. Smaller values (K = 3-6) are often preferred as they can capture local patterns in the data more effectively, which is particularly useful in scenarios where emotions are characterized by subtle and localized changes in brain activity. However, smaller values of K may also increase the sensitivity of the model to noise and outliers in the data (Kataria & Singh, 2013). On the other hand, larger values (K = 7-10) may lead to smoother decision boundaries and provide more robustness to noise in the data. This can be advantageous when dealing with EEG datasets that contain a high degree of variability or artifacts. However, larger values of K can also result in the model overlooking finer details in the data and potentially missing important patterns (Kataria & Singh, 2013). The value for K is mostly chosen to be uneven to avoid ties between the nearest neighbors values (Adeniyi, Wei, & Yongquan, 2016)

EEG data often exhibits intricate, non-linear relationships between recorded features and labels. KNN's inherent ability to capture non-linear relationships without imposing strict assumptions about data distribution renders it highly suitable for modeling the nuanced patterns present in EEG signals (Kataria & Singh, 2013). Furthermore, EEG recordings are inherently noisy due to various artifacts and other physiological interferences. Examples of commonly found artifacts are eye blinks, muscle activity, and electrode drift. KNN's robustness to noise allows it to effectively classify emotions in the presence of such disturbances, as it relies on multiple neighboring data points rather than being influenced by individual noisy observations. KNN also offers scalability and flexibility, which is beneficial when dealing with large datasets (Kataria & Singh, 2013). When considering modern EEG equipment, capturing data from 128 to 256 electrodes, KNN can accommodate the high-dimensional nature of these datasets and scale efficiently to process large volumes of data (Kataria & Singh, 2013). Furthermore, it is simple to implement, as it does not depend on complex parameter tuning or training procedures. This characteristic facilitates rapid prototyping and experimentation, enabling researchers to quickly iterate and refine their models based on performance feedback (Kataria & Singh, 2013).

Various past classification studies have been conducted using the KNN model on large datasets. An example, besides the study by Amin et al. (2017), is a study by Bhattacharyya, Khasnobish, Chatterjee, Konar, and Tibarewala (2010), where KNN was used to classify raw EEG data into left and right limb movements. The study compared KNN's performance to two other models, namely linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA). Two classification approaches were taken for all three models: one where all features were considered individually, and another where all features were combined. KNN achieved an accuracy of 75.71% when considering all features together. Additionally, KNN demonstrated especially high accuracies when applied to specific feature vectors derived from EEG signals. The average band power estimate vector showed the highest accuracy (84.29%) with the KNN algorithm. The paper's results highlight KNN's effectiveness when applied to large datasets, specifically concerning simple classification between left en right limb movement. The studies by Amin et al. (2017) and Bhattacharyya et al. (2010) indicate KNN's effectiveness in reliably classifying EEG signals during cognitive tasks with a high degree of accuracy. This study further corroborates the efficacy of the KNN model in handling large datasets and reliably classifying EEG signals during complex cognitive tasks.

2.5 *Classification with SVM*

The Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification tasks (Fletcher, 2009). SVM works by identifying the optimal hyperplane that separates different classes in the feature space. The hyperplane is determined by support vectors, which are the data points closest to the decision boundary. The key idea behind SVM is to maximize the margin between the hyperplane and the nearest data points of each class. This margin ensures a robust separation between classes. Once the hyperplane is determined, SVM assigns new data points to one of the classes based on which side of the hyperplane they fall. Overall, SVM operates by finding the optimal hyperplane that maximizes the margin between classes, making it a powerful algorithm for classification tasks (Fletcher, 2009).

As previously mentioned, EEG data is often non-linear, with intricate relations between features and labels. In cases where the data is not linearly separable, SVM uses a technique called kernel trick to map the input features into a higher-dimensional space where the data becomes linearly separable. This allows SVM to construct complex decision boundaries that can effectively separate different classes, even in non-linearly separable datasets (Fletcher, 2009). SVM is inherently a binary classifier, meaning it separates data into two classes. However, the One-vs-Rest (OvR) strategy is a classification approach used with SVMs to handle multi-class classification problems (Hong & Cho, 2006). In OvR, for each class in the dataset, a separate binary classifier is trained to distinguish that class from all other classes combined. During prediction, the class with the highest decision function value from the binary classifiers is assigned to the input sample. By training multiple binary classifiers, each specialized in distinguishing one class from the rest, OvR leverages the inherent discriminative power of SVMs to accurately predict multiple classes. The binary nature of the classifiers enables them to capture complex decision boundaries and extract relevant features for each class, resulting in robust and accurate predictions across multiple classes (Hong & Cho, 2006).

Utilizing the OvR strategy with an SVM classification model is therefore a suitable method of classifying EEG data into 15 emotional labels (Hong & Cho, 2006). In addition to its effectiveness in classification tasks, SVMs offer scalability, making them well-suited for handling large EEG datasets with numerous samples and features from multiple electrodes (Fletcher, 2009). SVMs also demonstrate strong generalization capabilities, crucial for accurately classifying emotions from EEG signals. Their ability to learn from training data and make accurate predictions on unseen data, increases the likelihood of reliable performance in real-world scenarios (Fletcher, 2009). Furthermore, SVMs incorporate regularization techniques to prevent overfitting, enhancing their ability to generalize to new instances of EEG signals. By controlling the complexity of the decision boundary and mitigating the influence of noise, SVMs provide robust and accurate classification of emotions from EEG data (Fletcher, 2009; Wang & Wang, 2021).

Because of these strengths, SVMs have been used for many classification problems based on large data with complex relations. A study by Subasi and Gursoy (2010) aimed to develop a comprehensive signal processing and analysis framework for EEG data, specifically focusing on the detection of epileptic seizures. The proposed framework employed Discrete Wavelet Transform (DWT) to decompose EEG signals into frequency sub-bands, followed by the extraction of statistical features to represent the distribution of wavelet coefficients. Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Linear Discriminant Analysis (LDA) were utilized to reduce the dimensionality of the data. Subsequently, SVM classifiers were trained using these features to distinguish between epileptic seizure and non-epileptic states. The SVM wielded accuracies of approximately 98% (PCA), 99% (ICA), and 100% (LDA). SVM was particularly useful for this research due to its ability to construct complex decision boundaries in high-dimensional feature spaces, making it well-suited for discriminating between epileptic seizure and non-seizure states based on EEG data with diverse patterns and characteristics.

2.6 *KNN versus SVM*

Given the complex and extensive dataset that is used in the current research, for **SQ1**, it is hypothesized that the SVM, utilizing the OvR strategy, will yield a higher classification accuracy compared to KNN. The complex and noisy nature of EEG recordings, characterized by artifacts like eye blinks and muscle activity, presents challenges for classification tasks. However, the SVM model excels in handling large, complex EEG datasets due to its ability to construct optimal hyperplanes that maximize the margin between classes, even in non-linear data. This method scales well to large datasets with many features and data points, while preventing overfitting with regularization techniques.

Despite KNN's mentioned strengths, past research by Amin et al. (2017) and (Wang & Wang, 2021) which included accuracy scores for both machine learning models, each reported higher accuracies using the SVM model. SVM's ability to construct intricate decision boundaries in high-dimensional feature spaces and its capacity to handle nonlinear relationships through the kernel trick offer it a unique advantage. Furthermore, SVM's One-vs-Rest (OvR) strategy allows it to handle multi-class classification tasks effectively by training multiple binary classifiers specialized in distinguishing one class from the rest, resulting in robust and accurate predictions across multiple classes (Hong & Cho, 2006). Overall, it is anticipated that SVM will outperform KNN in accurately classifying the diverse emotional states present in the dataset provided by (Onton & Makeig, 2022).

H1: ***"The SVM model will yield a higher classification accuracy compared to the KNN model."***

2.7 *Classification & Unique Data*

The priorly discussed study by Hsu et al. (2022), which makes use of the same dataset as the current study, utilized an unsupervised multimodel AMICA decomposition approach to investigate EEG dynamics during emotion imagination. Its objectives included identifying dominant EEG models, exploring their relationship with different emotions, analyzing spatio-temporal EEG dynamics, and uncovering active neurophysiological sources involved with human emotion. The findings showed consistent EEG segment separation during emotion periods, with significant differences in dipole density observed in brain regions like the prefrontal cortex, cingulate cortex, insula, motor cortex, and visual cortex. Moreover, the study revealed variability in emotional processes across different emotions, however, these were inconsistent among participants.

The research indicates that patterns of brain activity linked to specific emotions tend to be more consistent within individuals compared to across different individuals experiencing the same emotion (Hsu et al., 2022). This suggests that even when individuals experience the same emotion, variations exist in the neural processes and brain regions involved in generating those emotions (Hsu et al., 2022). EEG is capable of reflecting brain activity in real-time, through the use of it Hsu et al. (2022) discovered similarities and differences between participants' brain patterns when feeling an emotion.

One key finding by Hsu et al. (2022) is the identification of distinct EEG dynamics associated with specific emotional states. Individuals experiencing emotions such as happiness, sadness, or fear exhibit discernible patterns of neural activity that can be reliably differentiated using advanced signal processing techniques. These distinct EEG patterns provide valuable insights into what parts of the brain are involved when with diverse emotional experiences. However, despite the presence of distinct EEG patterns, similarities in brain activity across individuals experiencing the same emotion were also discovered (Hsu et al., 2022). Despite variations in individual neuroanatomy and life experiences, there exist similarities in the neural networks activated in participants during emotional states. This suggests the existence of a shared neural "signature" for certain emotions, underscoring the robustness and universality of emotional processing mechanisms. While there are these signatures, there is also considerable variability among individuals. Factors such as genetic predispositions, experiences, and cultural background contribute to this person-to-person variability in neural responses to emotions (Hsu et al., 2022). Consequently, while overarching patterns may exist, the precise neural structure of emotional experiences can vary significantly from person to person.

When training a machine learning model, the train and test splits typically consist of a random selection of respectively 80 and 20 percent of the total dataset (learn Developers, 2024). This random allocation does not take participant numbers or any other variables into account. As mentioned before, brain activity attributed to specific emotions tends to be more consistent within individuals when feeling a different emotion, than across different individuals experiencing the same emotion (Hsu et al., 2022). Thus, when researching emotional classification there could be a significant increase or decrease in classification accuracy when data from participants is included in both the test, and train dataset. Accuracy could increase since the algorithm recognizes the participants' brain activity patterns, therefore, being able to accurately classify it. Or the accuracy could decrease since the model may become overly sensitive to the unique characteristics of the participants' brain activity included in both sets, leading to reduced generalizability across all participants. Investigating this aspect of dividing train and test data could assist future research in comprehending the consequences of incorporating the same participant in both data splits. When considering the outcome to **SQ2**, even though generalizability will be reduced, it is hypothesized that the accuracy of both the KNN and the SVM model will increase when unique data points from multiple participants are included in the train and test due to the recognition of brain activity patterns. If true, future researchers can consider this information when deciding how to train-test split their data.

H2: ***"The accuracy of the KNN and SVM will increase when the test split exclusively includes data points derived from unique emotions of participants whose data on different emotions is included in the train split, when compared to the data being randomly divided between the two splits."***

2.8 *Emotions & the Brain*

The brain is central to human cognition, communication, and learning (Liu, Zhang, Li, & Kong, 2021). In the context of EEG emotion recognition, understanding the roles of major brain regions is crucial (Liu et al., 2021). The brain consists of three primary divisions: the brain stem, cerebellum, and cerebral cortex. Among these, the cerebral cortex is particularly noteworthy since it is responsible for higher cognitive functions, such as thinking and emotion (Liu et al., 2021). Within the cerebral cortex, as depicted in Figure 1, we find the parietal, frontal, temporal, and occipital lobes, each with distinct roles.

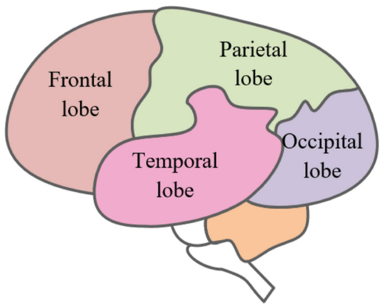

Figure 1: Brain map depicting the four lobes of the cerebral cortex. The image is provided by Liu et al. (2021).

Extensive research has been conducted on brain regions involved with emotion (Heilman, Gilmore, Li, & Kong, 1998; Pessoa, 2017; Šimić et al., 2021). Pessoa (2017) suggests that only a subset of brain structures are involved with emotion, namely: amygdala, periaqueductal gray, hypothalamus, ventral striatum, orbitofrontal cortex, insula, and medial PFC. According to Heilman et al. (1998), dysfunction of the cerebral cortex is associated with disorders of emotional communication. Specifically, dysfunction in any part of the cerebral cortex (posterior and anterior portions of the neocortex, and the left and right hemisphere) effects the communication of emotions. These findings suggest that the communication of emotions involves the cerebral cortex in its entirety. However, Šimić et al. (2021) suggests that when feeling emotions, humans do not use all brain structures. The parts that are involved are the insula, ventromedial prefrontal, anterior cingulate, amygdala, putamen, ventral tegmental area, ventral striatum, and caudate nucleus (Šimić et al., 2021). Additionally, Hsu et al. (2022) identified that when participants experienced emotion there were significant differences in brain activity in the prefrontal cortex, cingulate cortex, insula, motor cortex, and visual cortex.

Little research has been done on the change in accuracy when a specific subset of the recorded electrodes is trained and tested on. A study by Zhang and Chen (2020) investigated emotion recognition accuracies when considering different subsets of electrodes using the DEAP dataset. The study suggests the involvement of specific brain areas in emotion generation, and demonstrated that using a select few EEG electrodes placed on the frontal and central area of the scalp can yield accurate classification results of approximately 80%. EEG primarily measures electrical activity generated by large populations of neurons in the cerebral cortex, which is the outer layer of the brain. It can also detect signals from deeper brain structures to some extent. However, these signals are typically weakened by the surrounding tissue and skull, so they are not as prominent in EEG recordings compared to the activity from cerebral cortex. Thus, to answer **SQ3**, a subset exclusively consisting of electrodes placed on brain structures involved with emotion in the cerebral cortex was created. Pessoa (2017) and Šimić et al. (2021) mention multiple brain regions involved with feeling emotion, from these regions only some are located in the cerebral cortex. Namely, the prefrontal cortex, the anterior cingulate cortex, the insular cortex, the orbitofrontal cortex, and neural groups in the temporal lobe. Electrodes placed on those regions are included in the subset that will be presented in the Methods section 3.3.

Based on the papers by Pessoa (2017), Šimić et al. (2021), Hsu et al. (2022), and Zhang and Chen (2020) this study hypothesizes that the accuracy of the SVM and KNN models will not be significantly affected even when large amount of data are omitted, since all relevant brain structures' data is captured and included. The omittance of this irrelevant data could theoretically even lead to an increase in classification accuracy, since only relevant data affects the models' predictions. However, since the subset of electrodes presented in this paper only takes into account the placements of electrodes on the cerebral cortex, valuable data from relevant deeper brain structures might be lost. This might lead to a decrease in classification accuracy for both the SVM and KNN model. Since EEG primarily measures electrical activity generated in the cerebral cortex, and a slight negative and positive impact on the models' performance is expected, the overall accuracy of output will most likely not be affected. If true, this could suggest that the selected subset of electrodes is most relevant for any future research into emotion classification. This information can help future researchers greatly reduce monetary costs by reducing the need for expensive equipment and lowering computational cost.

H3: ***"The accuracy of the SVM and KNN models will not be affected by the exclusion of electrodes that do not detect signals directly originating from brain structures involved with emotion."***

3 METHODS

3.1 *Experimental Set-up*

The dataset used for this study is publicly available at the OpenNeuro website (`https://openneuro.org`) through the following accession number: 'ds003004'. It was fully collected and preprocessed by Onton and Makeig (2022). The dataset consists of the recordings of 34 participants during a self-paced emotion imagination task. The 34 volunteer participants consisted of 19 (55.88%) females and 15 (44.12%) males, with a mean age of 25.5. The oldest participant was 38 and the youngest participant was 18 years old. All participants were healthy, with reported normal cognitive conditions. Before starting the experiment, all participants signed an informed consent letter which met the UCSD institutional review board requirements (Hsu et al., 2022).

The experiment aimed to induce different emotions in the participants through the use of recorded voice narratives. A total of 15 emotional states were induced for each participant using this method. These emotions were classified as either positive or negative. The positive emotions included: awe, joy, happiness, love, compassion, contentment, relief, and excitement. The negative emotions included: frustration, anger, sadness, fear, jealousy, grief, and disgust. The experiment was divided into 30 event codes, which included: an introduction clip, EEG baseline activity recording, instruction clips, and sound clips that describe a scenario for the target emotion (Onton & Makeig, 2022). These event descriptors make clear that each session commenced with 2 minutes of closed-eyed silent rest, followed up by clips of general instructions. Thereafter, the participant followed a 5-minute guided induced relaxation to foster an inwardly focused state of mind. In the subsequent events, participants were presented with a series of voice-guided recordings instructing them to recall or imagine scenarios designed to evoke a vivid, embodied experience of the target emotion. All scripts of the voice-guided recordings are provided in JSON files included with the dataset. Two examples of these scripts provided by (Onton & Makeig, 2022, folder subjects) are:

> *FRUSTRATION (negative): "Something is not going as you would wish. You are beginning to feel agitated. Your body feels tense and uncomfortable, yet you feel helpless to change your situation. Perhaps your computer has crashed, deleting valuable files, or your ride to the airport is now an hour late. Or perhaps you are in a hurry but stuck in traffic. Gradually, you are overwhelmed with a profound and utter FRUSTRATION. Let this feeling of frustration and annoyance affect your whole body."*
>
> *"EXCITEMENT (positive): "You are beginning to feel quick, energetic, uplifted. Perhaps you sense this is really your lucky day. Perhaps something you dreamed of experiencing is finally about to arrive, something that opens up new, exciting possibilities for you. You are finally getting your chance to experience something you have long been looking forward to. Your body is filled with EXCITEMENT that bubbles up and out of you in every direction."*

Participants were instructed beforehand to take their time in recalling a past or imagining a possible scenario that would evoke the target emotion authentically. They were not given any time constraints or signals to indicate when to start or stop. Instead, participants were asked to press a button with their right hand when they began to feel the targeted emotion, and another button with their left hand when the feeling started to fade. This prompted a set of instructions lasting 40 seconds to relax and reset. After that, a 10-second audio clip prepared them for the next emotion imagination period. This is how it proceeded for all 15 target emotions, which were ordered the same for all participants. The experiment concluded with a post-baseline EEG recording (Onton & Makeig, 2022). No data was collected regarding the mental imagery participants recalled or imagined during the voice-guided experiment.

Due to the use of the handheld button to indicate the effect of an emotion, minimal movement was required from participants to complete this task, and the EEG-data could be recorded with relatively little noise. The experiment yielded a dataset consisting of approximate 3 to 5 minute recordings for each of the emotions per participant. The equipment used for recording the EEG data used 256 scalp channels with a Biosemi ActiveTwo system, positioned on the scalp with the standard 10-20 system as shown in Figure 2 (Hsu et al., 2022). It used a sampling rate of 256 Hz per channel with 24-bit resolution, which provides classification models with rich data for all 15 emotions through the high pool of channels (Hsu et al., 2022).

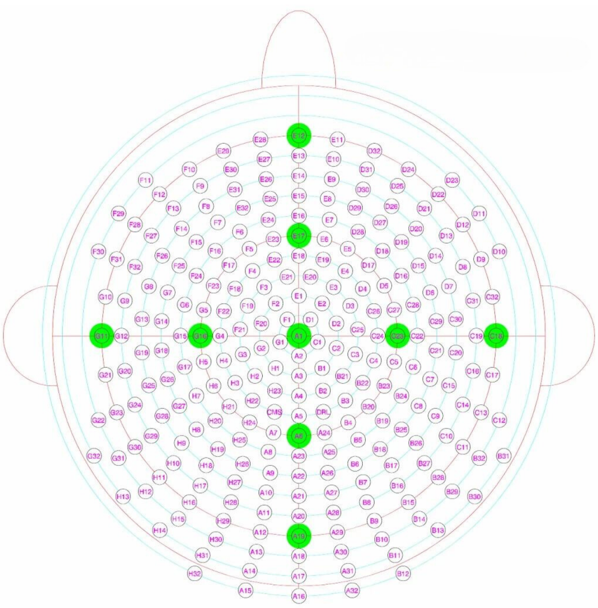

Figure 2: Visual representation of 256 channels positioned on the scalp utilizing the 10-20 system. The image is provided by Cortech Solutions (Accessed: 2024-05-31)

3.2 *Preprocessing & Dataset Description*

The dataset has been preprocessed by Hsu et al. (2022) in collaboration with the conductors of the experiments. The preprocessing consisted of five steps. The first preprocessing step entailed the removal of all channels of which the electrodes were poorly connected to the skin. In case a channel contained extremely abnormal activity patterns, researchers assumed that there was interference in the channel recording because of poor skin contact (Hsu et al., 2022). Secondly, the data was passed through a high pass filter of 1 Hz. The filter allowed for the removal of low-frequency fluctuations, which could have otherwise obscured the underlying neural activity. The data was re-referenced using Common Average Reference (CAR). CAR involves subtracting the average signal across all electrodes from each individual electrode's signal. This process helps to remove noise sources that affect all electrodes equally, such as electrical interference or environmental noise, while preserving the original EEG signal Hsu et al. (2022). Then Artifact Subspace Reconstruction was used to automatically remove artifacts from the data. Artifacts refer to any unwanted signals in the EEG recordings that are not generated by the brain's electrical activity such as eye blinks. Lastly, some channels were filtered out of the dataset, byselecting a subset of channels from the original EEG recording and discarding the rest. In this context it was used because the researchers aimed to address practical issues such as computational resources, processing time, and variations across participants. The variations across participants are especially important since the previous preprocessing steps left the dataset with missing channels for some participants. After preprocessing, the dataset was left with the recordings from 34 participants across 128 channels of scalp electrodes that provided the most representative coverage of the scalp while ensuring even spacing between channels Hsu et al. (2022).

3.3 *Data Analysis & Software*

To answer the research questions, two machine learning models were each trained and tested three different times with different train and test data. Visual Studio Code (Python) was used for all parts of the data analysis in this paper. Firstly, to answer the main research question (**RQ**), the Support Vector Machines and K-Nearest Neighbor models were trained and tested on the preprocessed dataset. The SVM and KKN machine learning models were chosen because of the high classification scores they yielded in past research, after being trained on the DEAP dataset (Wang & Wang, 2021; Yu & Wang, 2022). Their ability to handle non-linearity, robustness to noise, and provision of interpretable results make SVM and KNN well-suited for emotion classification using EEG data.

To judge the models' performance, their classification accuracy across all emotions were calculated. To answer **SQ1**, a standard train/test split with the ratio 80/20 was used to divide the data and the accuracies of the KNN and SVM models were thereafter compared to each other. As mentioned before, research has shown that patterns of brain activity associated with different emotions are more consistent within individuals than between different individuals experiencing the same emotion (Hsu et al., 2022). To answer **SQ2**, an untrained version of both models was trained and tested on a different test and train split. The train/test split was still in ratio 80/20, however, the test set exclusively contained data from emotions of participants who had data on other emotions in the training set. The train and test data did not contain data from the same label and participant, merely data from the same participant while feeling different emotions. The accuracies of the KNN and SVM models were reported and compared to the accuracies of **SQ1**.

A subset of electrodes was selected to answer **SQ3**. The subset consists of 25 out of the 128 preprocessed electrodes, namely: Fp1, Fp2, AF3, AF4, F7, F8, F3, Fz, F4, FC5, FC1, FC2, FC6, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, C3, Cz, C4. This selection is based on the previously mentioned (section 2.8) research by Zhang and Chen (2020), Pessoa (2017), and Šimić et al. (2021). Pessoa (2017) and Šimić et al. (2021). Additionally, research from Zhang and Chen (2020) on emotion classification divided the 32 electrodes in the DEAP dataset into 13 partially overlapping subsets of electrodes. Results showed a high contribution to emotion classification accuracy from the 25 electrodes named above (Zhang & Chen, 2020). A standard train/test split of 80/20 was used to train and test the data. The accuracies that the SVM and KNN models yielded were thereafter compared to the accuracies yielded from the **SQ1**.

the Wilcoxon Signed-Rank Test was selected to evaluate significance difference across the accuracy scores of the 15 emotion classes to answer all three subquestions. This test is specifically designed for paired data comparison, allowing for the assessment of changes in accuracy scores across different experimental conditions (Laerd, Accessed 2024-05-31). The Wilcoxon Signed-Rank Test does not assume normality and is robust to outliers, making it suitable for analyzing paired data obtained from the same participants to test different conditions. Additionally, to address concerns regarding multiple comparisons across the three sub-questions, the Bonferroni correction method was applied. This correction adjusts the significance level (α) to mitigate the increased risk of Type I errors associated with conducting multiple hypothesis tests using the following formula:

$$\alpha_{adjusted} = \frac{\alpha}{n}$$

By utilizing these statistical test and corrections, the study aims to provide robust and reliable insights into the classification accuracy of SVM and KNN models in emotion recognition tasks.

4 RESULTS

4.1 *Main Research Question & Sub Question 1*

All results will subsequently be presented in their own section. Each section will start out with the overall accuracies that the models yielded across the classification of the 15 emotions for that sub-question. Followed up by the results of the corresponding Wilcoxon Signed-Rank Test. Firstly, to answer the **RQ** and **SQ1**, the SVM model using the One-vs-Rest strategy and the KNN model were trained and tested on the entire preprocessed dataset. The overall accuracies achieved by both models when classifying all 15 emotions are presented is Table 1.

Table 1: Classification Accuracies for **SQ1**

Model	**Accuracy**
KNN	0.847
SVM (OvR)	0.865

The accuracies across the 15 emotions of both models were compared using the Wilcoxon Signed-Rank Test, resulting in the following p-value:

$$p\text{-value} = 0.0282$$

To address concerns regarding multiple comparisons across the three sub-questions, the Bonferroni correction was applied to correct the significance level (α) accordingly:

$$\tfrac{0.05}{3} = 0.0167$$

In sum, the p-value exceeds the significance level. Thus, the numerical difference in accuracy between the KNN and SVM model is not significant.

4.2 *Sub Question 2*

To answer **SQ2**, the SVM model using the One-vs-Rest strategy and the KNN model were trained and tested on the entire preprocessed dataset. The train and test splits of data were still kept in the ratio 80/20. However, the data in the test set solely consisted of data from participants in the train set feeling different emotions. The overall accuracies achieved by both models using this data split are presented is Table 2.

Table 2: Classification Accuracies **SQ2**

Model	**Accuracy**
KNN	0.861
SVM (OvR)	0.876

The accuracies across the 15 classes of the trained KNN models and SVM models from **SQ1** and **SQ2** were compared using the Wilcoxon Signed-Rank Test, with the following resulting p-values:

$$\text{KNN } p\text{-value} = 0.108$$

$$\text{SVM } p\text{-value} = 0.184$$

Both resulting p-values exceed the significance level of 0.0167. Thus, the numerical difference in accuracy between the KNN and SVM model is not significant.

4.3 *Sub Question 3*

To answer **SQ3**, the SVM model using the One-vs-Rest strategy and the KNN model were trained and tested on a subset of the dataset used to investigate **SQ1**. The subset used in this study is mentioned in section 3.3 of this paper. The overall accuracies achieved by both models using this data split are presented is Table 3.

Table 3: Classification Accuracies for **SQ3**

Model	**Accuracy**
KNN	0.771
SVM (OvR)	0.826

The accuracies of the trained KNN models and the accuracies of the SVM models from **SQ1** and **SQ3** were compared using the Wilcoxon Signed-Rank Test, with the following resulting p-values:

$$\text{KNN } p\text{-value} = 0.046$$

$$\text{SVM } p\text{-value} = 0.191$$

Both resulting p-values exceed the significance level of 0.0167. Thus, the difference in performance between the models from **SQ1** and **SQ3** is not significant.

5 DISCUSSION & CONCLUSION

5.1 *Hypotheses & Results*

This study aims to add to the current scientific literature on EEG-based emotion classification. To accomplish this, it investigated a main research question and three sub-questions. This paper hypothesized the following:

H1: ***"The SVM model will yield a higher classification accuracy compared to the KNN model."***

As can be seen in Table 1, the accuracies of the KNN and SVM model were about 85% and 87% respectively. To answer **SQ1**, the accuracies of the two models were compared using the Mann-Whitney U Test, which resulted in a p-value of 0.0282. After applying the Bonferroni correction to account for multiple comparisons, the adjusted significance level became 0.0167. Due to the Bonferroni correction, the resulting p-value is larger than the significance level, which indicates that the observed difference in classification accuracies between the SVM and KNN models is not statistically significant. This means that the findings in this study do not support **H1**.

H2: ***"The accuracy of the KNN and SVM will increase when the test split exclusively includes data points derived from unique emotions of participants whose data on different emotions is included in the train split, when compared to the data being randomly divided between the two splits."***

As can be seen in Table 2, The KNN model yielded an accuracy of about 86% and the SVM model yielded an accuracy of about 88%. The trained models in **SQ2** outperformed the models in **SQ1**. Two Wilcoxon Signed-Rank Tests were performed to test for significance difference between the tested KNN model of **SQ1** and **SQ2**, and between the tested SVM model from **SQ1** and **SQ2**. For the KNN models, the resulting p-value was 0.108, and for the SVM models, it was 0.184. Given a significance level of 0.0167, both p-values exceed this threshold. Therefore, no significant difference in accuracy was found between the models, which suggest that the findings of this study do not support **H2**.

H3: ***"The accuracy of the SVM and KNN models will not be affected by the exclusion of electrodes that do not detect signals directly originating from brain structures involved with emotion."***

Table 3 shows the accuracies yielded by the KNN and SVM models while trained on data from only a subset of electrodes. The accuracies were approximately 77% and 83% respectively, which are both numerically lower than the accuracies of each model in **SQ1**. To test for significance, two Wilcoxon Signed-Rank Tests were performed between the accuracies of **SQ1** and **SQ3**. These tests resulted in p-values of 0.046 and 0.191 on the KNN and SVM models, respectively. Both p-values are above the significance level of 0.0167, which indicates a non-significant difference between the performance of both the KNN and SVM models between **SQ1** and **SQ3**. Since the performance of both models is not significantly different when trained on a subset of data, this study's findings support **H3**.

RQ: ***"How accurately can Support Vector Machines and K-Nearest Neighbor models classify 15 different emotions felt across multiple minutes from electroencephalographic data?"***

Table 1 contains the results of the KNN and SVM models needed to answer the main **RQ**. the KNN model yielded an accuracy of 0.798, while SVM yielded 0.835. This entails that the SVM and KNN models can classify 15 different emotions felt across multiple minutes from electroencephalographic data, with respective accuracies of about 87% and 85%.

5.2 *Past Literature & Societal Implications*

The findings of this study contribute to the growing body of research on emotion classification using EEG data, aligning with and extending insights from past studies. In particular, this research supports the results of Wang and Wang (2021) and Yu and Wang (2022), who reported high accuracies for KNN and SVM models in emotion classification using the DEAP, SEED, and DREAMER datasets. The current study's conventionally trained (**SQ1**) SVM model achieving an accuracy of 87% and KNN model achieving 85% accuracy are consistent with the range of 50-97% reported in previous literature. While the SVM model outperformed the KNN model in this study, the difference was not statistically significant (p-value of 0.0282). However, The adjusted significance level of the standard 0.05 to 0.0167 with the Bonferroni correction is known to be of the more extreme corrections to account for multiple comparisons. If a different correction method was applied, the findings in this paper might have more closely resembled the findings of significant difference between the SVM and KNN models in past research (Amin et al., 2017; Subasi & Gursoy, 2010). The results of this paper, though not conclusively favoring SVM, still align with findings from Amin et al. (2017) and Subasi and Gursoy (2010), where SVM higher accuracies than KNN in classifying EEG data. Specifically, Amin et al. (2017) reported SVM reaching 99.11% accuracy on low-frequency coefficients, while Subasi and Gursoy (2010) noted SVM's superior performance in detecting epileptic seizures.

This study hypothesized that including unique data points from multiple participants in both the train and test datasets would increase model accuracy. This hypothesis was in line with past research by Hsu et al. (2022) which underlined the consistency in patterns of brain activity linked to specific emotions, which tend to be more consistent within individuals compared to different individuals experiencing the same emotion. However, the results of this paper contrast with past findings by not finding a significant difference in performance between **SQ1** and **SQ2**. By finding that including multiple participants' data does not significantly enhance model accuracy, the findings of this study underscore the complexity of EEG-based emotion classification and the potential limitations of generalizing neural patterns across individuals. The consistent numerically higher accuracy of the SVM model in both **SQ1** and **SQ2** further suggests its robustness in handling EEG data, particularly when inter-individual variability is a factor. Future research might benefit from exploring alternative strategies for data splitting and model training to better capture the nuances of emotional processing in diverse populations.

Based on past research by Pessoa (2017), Šimić et al. (2021), Zhang and Chen (2020) a subset of electrodes was created that contains only electrodes placed on brain regions involved with emotion. This study's results support these findings by demonstrating that both KNN and SVM models can effectively classify emotions based on EEG data, despite being provided with a subset of electrodes (provided in section 3.3). The results corroborate the importance of these brain regions in emotion processing, as suggested by Pessoa (2017) and Šimić et al. (2021), and reinforce the notion that EEG can capture meaningful neural signatures of emotional states. The use of EEG for emotion classification holds significant promise for various applications, including affective computing, mental health diagnostics, and human-computer interaction. The findings suggest that SVM may offer a slight edge in performance compared to KNN even when presented with less data (the subset) due to its ability to handle complex decision boundaries and multi-class classification more effectively. However, the non-significant difference in performance between SVM and KNN highlights the need for further exploration into optimizing model parameters and incorporating additional features to enhance accuracy. This finding supports the hypothesis that relevant data from critical brain structures can provide sufficient information for accurate emotion classification. It aligns with Zhang and Chen (2020), who demonstrated that using electrodes on the frontal and central scalp areas yielded high classification accuracies.

The findings of this study have multiple practical and societal implications. Improved emotion classification from EEG data has the potential to enhance various aspects of human life. In mental health, more accurate emotion detection can lead to better diagnostic tools and personalized treatment plans, ultimately improving patient outcomes. The findings in this study provide insights into the electrodes that are important to focus on during EEG emotion related research, which can decrease computing costs. In the field of affective computing, advancements in emotion recognition can facilitate the development of more intuitive and responsive human-computer interactions, leading to more user-friendly technology interfaces. Furthermore, this research can aid in creating more socially aware artificial intelligence, enhancing the ability of AI systems to understand and respond to human emotions appropriately. In healthcare, robotic aids equipped with advanced emotion recognition capabilities could provide better support for patients, both emotionally and physically. Additionally, understanding emotions more comprehensively can improve social interactions by fostering empathy and effective communication. These implications underscore the importance of continued research in EEG-based emotion classification, aiming to translate these scientific advancements into practical, real-world benefits.

5.3 *Study Limitations & Future Directions*

It is important to recognize several limitations of this study. Firstly, the study sample consisted of only 34 participants, which may limit the generalizability of the findings to larger populations. Additionally, the study did not control for or collect data on the imagined narratives of the participants while evoking emotions, potentially overlooking important contextual factors that could influence the recorded EEG data. Furthermore, data collection occurred in a controlled laboratory setting, which may have implications for the ecological validity of the results. Real-world emotional experiences often occur in dynamic and uncontrolled environments, and the artificiality of the lab setting could affect the participants' emotional responses and subsequent EEG patterns. The sheer volume of data collected in this study also posed practical challenges. The extensive dataset required considerable computational resources and time for model training, especially for the SVM model. The SVM model took around 5 times as long to be trained and tested compared to the KNN model for all subquestions. This limitation not only increased the complexity and duration of the study but also raised concerns about the scalability and efficiency of the proposed machine learning approaches. Lastly, this research made use of the Bonferroni correction method, which lowered the significance level from 0.05 to 0.0167. This correction is stricter than the corrections used by other research discussed in this paper (Wang & Wang, 2021; Zhang & Chen, 2020). Despite these limitations, the study contributes valuable insights into the classification of emotions from EEG data using machine learning techniques.

Future research endeavors should address these limitations by employing larger and more diverse participant samples to enhance the generalizability of the findings. Additionally, incorporating measures to capture imagined narratives during emotion evocation could provide a richer context and improve the accuracy of emotion classification. Furthermore, exploring data collection methods that enhance ecological validity, such as using portable EEG devices in real-world settings, would make the results more applicable to naturalistic scenarios. Optimizing computational resources for more efficient model training is also crucial to ensure the scalability of the proposed approaches. Future studies can investigate additional alternative data splitting strategies and model training techniques that account for across individual variability in emotional processing. This could involve using advanced machine learning methods, such as transfer learning, to leverage shared neural patterns while accommodating individual differences. Incorporating multimodal data, such as combining EEG with physiological signals or behavioral data, could also enhance the robustness and accuracy of emotion classification models. Moreover, research should explore the potential of personalized emotion classification models tailored to individual neural signatures. This personalized approach could lead to more accurate and reliable emotion detection systems, benefiting applications in mental health diagnostics, affective computing, and human-computer interaction. Investigating the use of different significance correction methods and comparing accuracy between the 15 mentioned emotions might also provide a more nuanced understanding of the statistical differences between models and align the findings more closely with prior research. Based on the mentioned information, two interesting future research questions could be:

RQ1 *"What are the effects of using portable EEG devices in real-world settings versus controlled laboratory environments on the robustness and ecological validity of emotion classification models? "*

RQ2 *"What are the optimal strategies for integrating imagined narratives into EEG-based emotion classification models to enhance accuracy and contextual relevance in both controlled laboratory settings and real-world environments?"*

5.4 *Conclusion*

This paper aimed to aid in the creation of a generalizable EEG-based emotion classification model by investigating classification accuracies of the SVM and KNN model across three sub-questions. To facilitate this goal, the current study utilized a recently released EEG dataset, containing rich EEG data on 15 different emotions experienced over multiple minutes.The findings contribute to the growing body of literature on EEG-based emotion classification by providing insights into the performance of SVM and KNN in distinguishing between different emotions. Overall, the results indicate potential for accurately classifying emotions using EEG data. This study found that SVM and KNN can classify 15 emotions felt across multiple minutes from electroencephalographic data with accuracies of 87% and 85%, respectively. Additionally, the examination of sub-questions revealed the efficacy of both KNN and SVM models in classifying emotions from EEG data, with a non-significant numerical advantage for SVM. The results align with previous research, validating the importance of including electrodes placed on brain regions involved with emotion. Future exploration of novel methodologies and optimization techniques will further contribute to the development of a highly accurate generalized model for classifying human emotions. The practical and societal implications of this research are significant, with potential applications in mental health diagnostics, affective computing, human-computer interaction, and healthcare.

REFERENCES

Adeniyi, D., Wei, Z., & Yongquan, Y. (2016). Automated web usage data mining and recommendation system using k-nearest neighbor (knn) classification method. *Applied Computing and Informatics*, *12*(1), 90-108. Retrieved from https://www.sciencedirect.com/science/article/pii/S221083271400026X doi: https://doi.org/10.1016/j.aci.2014.10.001

Alarcao, S. M., & Fonseca, M. J. (2017). Emotions recognition using eeg signals: A survey. *IEEE transactions on affective computing*, *10*(3), 374–393.

Amin, H. U., Mumtaz, W., Subhani, A. R., Saad, M. N. M., & Malik, A. S. (2017). Classification of eeg signals based on pattern recognition approach. *Frontiers in Computational Neuroscience*, *11*. Retrieved from https://www.frontiersin.org/articles/10.3389/fncom.2017.00103 doi: 10.3389/fncom.2017.00103

Andersen, P. A., & Guerrero, L. K. (1996). Chapter 3 - principles of communication and emotion in social interaction. In P. A. Andersen & L. K. Guerrero (Eds.), *Handbook of communication and emotion* (p. 49-96). San Diego: Academic Press. Retrieved from https://www.sciencedirect.com/science/article/pii/B9780120577705500059 doi: https://doi.org/10.1016/B978-012057770-5/50005-9

Bhattacharyya, S., Khasnobish, A., Chatterjee, S., Konar, A., & Tibarewala, D. (2010). Performance analysis of lda, qda and knn algorithms in left-right limb movement classification from eeg data. In *2010 international conference on systems in medicine and biology* (p. 126-131). doi: 10.1109/ICSMB.2010.5735358

Cimtay, Y., & Ekmekcioglu, E. (2020). Investigating the use of pretrained convolutional neural network on cross-subject and cross-dataset eeg emotion recognition. *Sensors*, *20*(7), 2034.

Cortech Solutions. (Accessed: 2024-05-31). *Da-at-hcs256.* https://cortechsolutions.com/product/da-at-hcs256/.

Cui, H., Liu, A., Zhang, X., Chen, X., Wang, K., & Chen, X. (2020). Eeg-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network. *Knowledge-Based Systems*, *205*, 106243.

Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* John Murray. Retrieved from https://darwin-online.org.uk/content/frameset?viewtype=text&itemID=F1142&pageseq=1

Dewey, J. (1895). The theory of emotion. *Psychological review*, *2*(1), 13.

Duan, R.-N., Zhu, J.-Y., & Lu, B.-L. (2013). Differential entropy feature for eeg-based emotion classification. , 81–84.

Fletcher, T. (2009). Support vector machines explained. *Tutorial paper*, *1118*, 1-19.

Frijda, N. H. (2007). Laws. In *The laws of emotion* (p. 1-366). New York: Psychology Press. Retrieved from https://www.taylorfrancis.com/

books/mono/10.4324/9781315086071/laws-emotion-nico-frijda doi: https://doi.org/10.4324/9781315086071

Heilman, K., Gilmore, R., Li, Y., & Kong, X. (1998). Cortical influences in emotion. *J Clin Neurophysiol.*, *15*(5), 409-4023. doi: 10.1097/00004691-199809000-00005

Holm, S. K., Kaakinen, J. K., Forsström, S., & Surakka, V. (2021). Self-reported playing preferences resonate with emotion-related physiological reactions during playing and watching of first-person shooter videogames. *International Journal of Human-Computer Studies*, *155*, 102690.

Hong, J.-H., & Cho, S.-B. (2006). Multi-class cancer classification with ovr-support vector machines selected by naive bayes classifier. In *International conference on neural information processing* (p. 155-164).

Hsu, S.-H., Lin, Y., Onton, J., Jung, T.-P., & Makeig, S. (2022). Unsupervised learning of brain state dynamics during emotion imagination using high-density eeg. *NeuroImage*, *249*, 118873. Retrieved from https://www.sciencedirect.com/science/article/pii/S1053811922000039 doi: https://doi.org/10.1016/j.neuroimage.2022.118873

Kataria, A., & Singh, M. (2013). A review of data classification using k-nearest neighbour algorithm. *International Journal of Emerging Technology and Advanced Engineering*, *3*(6), 354-360. Retrieved from https://www.researchgate.net/publication/353306410_A_Review_of_Data_Classification_Using_K-Nearest_Neighbour_Algorithm

Katsigiannis, S., & Ramzan, N. (2017). Dreamer: A database for emotion recognition through eeg and ecg signals from wireless low-cost off-the-shelf devices. *IEEE journal of biomedical and health informatics*, *22*(1), 98–107.

Koelstra, S., Muehl, C., Soleymani, M., J.-S. Lee, T. E., A. Yazdani, Pun, T., Nijholt, A., & Patras, I. (2012). Deap: A database for emotion analysis using physiological signals. *EEE Transactions on Affective Computing*, *3*(1), 18-31. Retrieved from https://www.eecs.qmul.ac.uk/mmv/datasets/deap/

Kuppens, P., & Verduyn, P. (2017). Emotion dynamics. *Current Opinion in Psychology*, *17*, 22-26. Retrieved from https://www.sciencedirect.com/science/article/pii/S2352250X16302019 (Emotion) doi: https://doi.org/10.1016/j.copsyc.2017.06.004

Laerd, S. (Accessed 2024-05-31). *Wilcoxon signed-rank test using spss statistics.* https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spss-statistics.php.

learn Developers, S. (2024). *sklearn model selection train test split.* https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html. (Accessed: 2024-06-19)

Levine, L. J., Lench, H. C., & Safer, M. A. (2009). Functions of remembering and misremembering emotion. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(8), 1059–1075.

Liu, H., Zhang, Y., Li, Y., & Kong, X. (2021). Review on emotion recognition based on electroencephalography. *Comput. Neurosci*, *15*, 1-15. doi: 10.3389/fncom.2021.758212

Niedenthal, P. M., & Ric, F. (2017). *Psychology of emotion*. Psychology Press.

Onton, J., & Makeig, S. (2022). "imagined emotion study". doi: 10.18112/openneuro.ds003004.v1.1.1

Pessoa, L. (2017). A network model of the emotional brain. *Trends in Cognitive Sciences*, *21*(5), 357-371. Retrieved from https://www.sciencedirect.com/science/article/pii/S1364661317300360 doi: https://doi.org/10.1016/j.tics.2017.03.002

Placidi, G., Di Giamberardino, P., Petracca, A., Spezialetti, M., & Iacoviello, D. (2016). Classification of emotional signals from the deap dataset. In *International congress on neurotechnology, electronics and informatics* (Vol. 2, pp. 15–21).

Plutchik, R., & Kellerman, H. (2013). *Theories of emotion* (Vol. 1). Academic press.

Saxena, A., Khanna, A., & Gupta, D. (2020, 01). Emotion recognition and detection methods: A comprehensive survey. *Journal of Artificial Intelligence and Systems*, *2*, 53-79. doi: 10.33969/AIS.2020.21005

Scherer, K. R., et al. (2000). Psychological models of emotion. *The neuropsychology of emotion*, *137*(3), 137–162.

Siuly, S., Li, Y., & Zhang, Y. (2016). Electroencephalogram (eeg) and its background. In *Eeg signal analysis and classification: Techniques and applications* (pp. 3–21). Cham: Springer International Publishing. doi: 10.1007/978-3-319-47653-7_1

Subasi, A., & Gursoy, M. I. (2010). Eeg signal classification using pca, ica, lda and support vector machines. *Expert systems with applications*, *37*(12), 8659–8666.

Wang, J., & Wang, M. (2021). Review of the emotional feature extraction and classification using eeg signals. *Cognitive Robotics*, *1*, 29-40. Retrieved from https://www.sciencedirect.com/science/article/pii/S2667241321000033 doi: https://doi.org/10.1016/j.cogr.2021.04.001

Yu, C., & Wang, M. (2022). Survey of emotion recognition methods using eeg information. *Cognitive Robotics*, *2*, 132-146. Retrieved from https://www.sciencedirect.com/science/article/pii/S2667241322000118 doi: https://doi.org/10.1016/j.cogr.2022.06.001

Zhang, J., & Chen, P. (2020). Selection of optimal eeg electrodes for human emotion recognition. *IFAC-PapersOnLine*, *53*(2), 10229-10235. Retrieved from https://www.sciencedirect.com/science/article/pii/S2405896320335163 (21st IFAC World Congress) doi: https://doi.org/10.1016/j.ifacol.2020.12.2753

Zheng, W.-L., & Lu, B.-L. (2017). A multimodal approach to estimating vigilance using eeg and forehead eog. *Journal of neural engineering*, *14*(2), 026017.

Šimić, G., Tkalčić, M., Vukić, V., Mulc, D., Španić, E., Šagud, M., ... Hof, R. (2021). Understanding emotions: Origins and roles of the amygdala.

Biomolecules, *11*(6), 823. doi: 10.3390/biom11060823