AutoML-based EEG signal Analysis in Neuro-Marketing classification using Biclustering method

Francisco Nauber Bernardo Gois[a], Joao Alexandre Lobo Marques[a], Bruno Riccelli dos Santos Silva[b], Paulo Cortez[b], Victor Hugo Costa Albuquerque[b]

[a]Laboratory of Applied Neurosciences,University of Saint Joseph, Macau, China
[b]Department of Engineering of Teleinformatics, Federal University of Ceara, Fortaleza, Ceará Brazil

---

Abstract

Background: Consumer neuroscience analyzes individuals' preferences through the assessment of physiological data monitoring, considering brain activity or other bioinformation to assess purchase decisions. Traditional marketing tactics include customer surveys, product evaluations, and comments. Understanding consumer neurological responses when seeing an ad or testing a product is crucial for product or brand marketing and mass production.

New method: In this work, we propose using the bi-clustering method to reduce EEG noise and automatic machine learning to classify brain responses. We utilized a neuromarketing EEG dataset that contains EEG data from product evaluations from 25 participants, collected with a 14-channel Emotiv Epoch+ device while examining consumer items. The methodology included the Welch Transform for filtering the EEG raw data and training different classification models with the best-converted signal biclusterings. The H2O.ai AutoML library was then used to select the optimal biclustering and models.

Results: The proposed method achieved an F1-Score of 0.95, indicating a high level of accuracy in classifying consumer preferences based on EEG data. Two thresholds were used for evaluation: the first threshold indicating customer satisfaction and the second threshold indicating dissatisfaction, with values in between classified as uncertain.

Comparison with existing methods: Our approach outperformed the state of the art, which typically achieves lower accuracy rates. For instance, previous methods using traditional machine learning models like SVM and logistic regression have shown accuracy rates ranging from 51.48% to 70%. Our use of biclustering and AutoML significantly improves upon these results.

Conclusions: The integration of biclustering for noise reduction and AutoML for classification demonstrates a powerful method for analyzing EEG signals in neuromarketing. This approach not only enhances the accuracy of predicting consumer preferences but also provides a robust framework for future research in consumer neuroscience. The high F1-Score achieved confirms the effectiveness of the proposed method, making it a valuable tool for understanding consumer behavior in real-time marketing scenarios.

Keywords: EEG signal Analysis, Consumer Neuroscience, Classification, Biclustering, Auto Machine Learning
PACS: 87.85.E, 87.19.le, 07.05.Mh, 87.19.lo

Author Contributions

The authors of this study, "AutoML-based EEG Signal Analysis in Neuro-Marketing Classification using Biclustering Method," have contributed significantly in various capacities to the research and preparation of this manuscript. Their specific contributions are detailed as follows:

Francisco Nauber Bernardo Gois: Conceptualization, Methodology, Investigation, Data Curation, Writing - Original Draft.
Joao Alexandre Lobo Marques: Supervision, Project Administration, Writing - Review & Editing.
Bruno Riccelli dos Santos Silva: Investigation,Writing - Original Draft ,Data Curation, Formal Analysis, Software.
Paulo Cortez: Resources, Validation, Writing - Review & Editing.
Victor Hugo Costa Albuquerque: Supervision, Funding Acquisition, Writing - Review & Editing.

Each author has approved the final version of the manuscript and agrees to be accountable for all aspects of the work.

1. Introduction

The application of neurosciences approaches to understand consumer preferences is known as Consumer Neuroscience, also interchangeably defined as Neuromarketing. It investigates the brain activity linked to preference and intention to buy. Traditional marketing strategies take into account a posteriori user input in the form of surveys, product reviews, or commentaries; nevertheless, these strategies fall short of properly capturing or explaining the real-time decision-making process of consumers. Moreover, understanding the consumer physiological responses while viewing a product ad may become an essential role for product or brand marketing validation and large-scale production.

1.1. Historical Context

The concept of ”consumer neuroscience” originates from the belief that neuroscientific and psychological techniques can be effectively applied by consumer researchers to better understand the factors motivating consumers' purchasing decisions. In order to evaluate the quality of the service provided to customers, these methods will be used in addition to or in instead of the more common practice of soliciting feedback directly from customers. In the past 15 years, there has been an exponential rise in the number of scholarly articles devoted to this subject [1, 2].

Neuromarketing uses non-invasive brain signal recording techniques to directly measure the response of a customer's brain to the marketing stimuli, superseding the traditional survey methods. Functional Magnetic Resonance Imaging (fMRI), Electroencephalography (EEG), Magnetoencephalography (MEG), Positron Emission Tomography (PET), and functional Near-Infrared Spectroscopy (fNIRS) are some examples of brain monitoring devices used in Neuromarketing research [3]. These techniques can be used together with other biometric sensors, with the idea to analyze the consumer decision making process more effectively [4].

Due to the growing popularity of Electroencephalography (EEG), mobile, user-friendly, and low cost EEG devices are being introduced to the market. The Emotiv EEG is one of these gadgets that lately gained popularity. It has already undergone testing in a number of applications including language processing, brain-computer interfaces, neuromarketing, and the detection of the P-300 component, with the overall finding that it is capable of capturing gratifying research data [5].

1.2. Importance of EEG Signal Analysis

In this study, we propose the use of the bi-clustering method be used to reduce noise in EEG signals and automatic machine learning to classify the brain signals from a group of 25 people in order to evaluate product acceptance and preference. The proposed technique outperforms the current state of the art by obtaining a F1-Score value of 0.95.

The three key contributions of this paper are:

- A method for noise reduction in time series utilizing biclustering.
- An AutoML method for classifying brain signals.
- The use of Biclustering with an uncertainty threshold for the model evaluation and classification.

2. Related Work

Several studies have explored the application of EEG signal analysis in the context of neuromarketing. EEG analysis allows for a deeper understanding of consumers' brain responses to various marketing stimuli, providing valuable insights for predicting consumer preferences.

Yadava et al. [6] presented the dataset used in the current work, where EEG signals were collected as participants viewed products on a computer screen. Using filtering and transformation techniques such as the Savitzky-Golay filter and Discrete Wavelet Transform, the authors achieved up to 70

Alimardani and Kaba [7] explored the use of deep learning methods for prediction tasks in neuromarketing, comparing the eficacy of convolutional neural networks (CNN) with traditional machine learning models. The results showed an accuracy of up to 51.48

Alnuman et al. [8] evaluated various time and frequency domain features of EEG signals in different brain regions, using the SVM classifier to identify the best features correlating with consumer emotions. After applying filtering, ICA, PCA, and normalization, the authors achieved classification accuracy ranging from 60.71

Ullah et al. [9] proposed a predictive model for identifying "Likes" and "Dislikes" of e-commerce products. Using a user-independent testing procedure, they assessed the eficacy of several classifiers, including Artificial Neural Networks, Logistic Regression, Decision Trees, K-Nearest Neighbors, and SVM.

Golnar-Nik et al. [10] investigated the feasibility of using EEG power spectrum features to predict consumer preferences in decision-making. Comparing SVM with LDA, they achieved accuracies of 63

Ma et al. [11] studied how brand familiarity and product category influence the brand extension decision-making process, revealing the neurological mechanisms underlying this phenomenon using a conventional ERP paradigm to record EEG signals.

Ausin-Azofra et al. [12] utilized various neuromarketing measurement devices (eye-tracking, facial coding, electrodermal activity, and EEG) to measure consumer responses to advertising, comparing the effectiveness of these methods in predicting product choice.

Garczarek-Bak et al. [13] compared neuromarketing-related measurements, such as EEG, EDA, and eye-tracking, to determine the best predictor of product choice among young adults.

Hakim et al. [14] used multiple EEG measures and machine learning techniques to improve preference prediction based solely on self-reports. They achieved an accuracy of 68.5

Hsu et al. [15] investigated the impact of subliminal advertising on hotel selection by measuring brain activity while participants watched hotel advertisements. The findings indicated that subliminal smiley face stimuli significantly influenced participants' hotel choices.

In addition to these studies, other significant works in the field include:

- Kahani et al. [16]: This study used deep learning models to analyze EEG signals for emotion recognition in neuromarketing. They achieved notable improvements in accuracy compared to traditional machine learning methods.

- Vecchiato et al. [17]: This research investigated the neurophysiological responses of individuals to TV commercials using EEG and other biometric measures, providing insights into consumer engagement and preference.

- Liu et al. [18]: This study explored the use of multimodal approaches combining EEG with other physiological measures to enhance the prediction of consumer behavior in marketing research.

- Goshvarpour et al. [19]: This work focused on analyzing emotional responses to advertisements using EEG signals and machine learning techniques, contributing to the understanding of emotional engagement in marketing.

3. Background

3.1. Artificial Intelligence applied for Neuromarketing

Artificial Intelligence (AI) decision-making approaches have been increasingly employed to assist professionals across various domains, such as robotics, medicine, security, and marketing, providing advanced techniques for classification, regression, and clustering tasks to understand data behavior or trends. In marketing, AI models can analyze vast amounts of data to uncover patterns that are not easily detectable by humans, thereby enhancing the ability to predict consumer behavior and preferences.

The integration of AI into neuromarketing represents a significant advancement in understanding consumer behavior. By leveraging AI algorithms, marketers can gain insights into the subconscious preferences of consumers, which traditional methods like surveys and focus groups may fail to capture accurately. This deeper understanding allows for the creation of more targeted and effective marketing strategies, ultimately leading to increased consumer engagement and sales.

Due to the high failure rates of newly launched products, which range from 80% to 97%, there is a critical need for more reliable methods to understand market issues, consumer behavior, and needs. Neuromarketing addresses this gap by applying neuroscience theories and techniques to marketing, aiming to comprehend the underlying processes that drive consumer decisions. By combining behavioral theories with methods and theories from neuroscience and AI research, neuromarketing provides a more comprehensive view of consumer behavior.

AI techniques, such as machine learning and deep learning, have been instrumental in analyzing EEG signals for neuromarketing applications. Machine learning models, including Support Vector Machines (SVM), Random Forests (RF), and Neural Networks, have been used to classify EEG signals with varying degrees of success. Deep learning approaches, particularly Convolutional Neural Networks (CNNs), have shown promise in capturing complex patterns in EEG data, leading to improved accuracy in predicting consumer preferences [7].

Recent advancements in AI have introduced automated machine learning (AutoML) frameworks, such as H2O.ai, which streamline the process of model selection and optimization. AutoML systems can automatically preprocess data, select the best features, and train multiple models to identify the most effective one for a given task. This automation significantly reduces the time and expertise required to develop high-performing models, making advanced AI techniques accessible to a broader range of researchers and practitioners [20].

Furthermore, the development of AI-driven tools for neuromarketing extends beyond EEG signal analysis. For example, researchers have explored the use of eye-tracking data, facial expression analysis, and physiological measures like heart rate and skin conductance to gain a holistic understanding of consumer responses to marketing stimuli. These multimodal approaches, which combine data from various sources, provide a richer and more nuanced view of consumer behavior [4].

In addition to improving the accuracy of consumer preference predictions, AI techniques also offer the potential for real-time analysis. This capability is particularly valuable in dynamic environments, such as online shopping platforms or live marketing events, where immediate feedback on consumer reactions can inform adaptive marketing strategies. Real-time AI analysis can enhance the personalization of marketing efforts, leading to more engaging and relevant consumer experiences.

Overall, the application of AI in neuromarketing holds significant promise for transforming how marketers understand and influence consumer behavior. By integrating advanced AI techniques with neuroscientific methods, researchers can develop more effective marketing strategies that are grounded in a deeper understanding of the human brain and its responses to marketing stimuli.

3.2. Biclustering

Biclustering, also known as co-clustering, bidimensional clustering, or subspace clustering, is an advanced data analysis technique that simultaneously clusters rows and columns of a data matrix. This method was first proposed by Hartigan in 1972 [21] and has since been extensively applied in various fields, including bioinformatics, where it was popularized by Cheng and Church for analyzing high-throughput gene expression data [22]. Unlike traditional clustering methods, which only group one dimension of the data matrix, biclustering provides a more nuanced view by revealing local patterns within the data.

One of the primary motivations for using biclustering is its ability to identify subsets of rows and columns that exhibit coherent patterns. This capability is particularly valuable in high-dimensional datasets where traditional clustering methods may fail to detect such localized structures. For instance, in gene expression data, biclustering can identify genes that co-express under specific conditions, providing insights into the underlying biological processes.

Biclustering techniques can be broadly categorized based on their goals: some aim to find a single bicluster at a time, while others attempt to identify multiple biclusters simultaneously. Methods that find one bicluster at a time, such as those proposed by Cheng and Church, iteratively search for biclusters, mask them with random numbers, and repeat the process to uncover additional biclusters [22]. On the other hand, techniques like the plaid model introduced by Lazzeroni and Owen aim to identify multiple biclusters in a single run, modeling the data as a superposition of several biclusters [23].

The application of biclustering in neuromarketing involves analyzing EEG signal data to identify patterns of brain activity that correlate with consumer preferences. By simultaneously clustering time points (columns) and EEG channels (rows), researchers can detect specific intervals and brain regions that exhibit significant responses to marketing stimuli. This dual clustering approach enhances the interpretability of EEG data, making it easier to relate neural responses to consumer behavior.

Several biclustering algorithms have been developed, each with its strengths and weaknesses. The choice of algorithm depends on the nature of the data and the specific objectives of the analysis. Commonly used algorithms include the Cheng and Church algorithm, which focuses on minimizing the mean squared residue of the biclusters, and spectral biclustering, which uses eigenvalue decomposition of the data matrix to identify coherent patterns [24]. Spectral biclustering, in particular, has gained popularity due to its robustness and ability to handle large, sparse datasets.

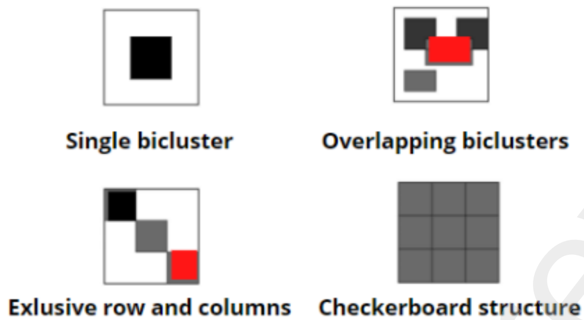
Advanced biclustering techniques have also incorporated machine learning methods to improve their performance. For example, Bayesian biclustering approaches use probabilistic models to account for uncertainty and noise in the data, providing more reliable bicluster identification [25]. Additionally, ensemble biclustering methods combine multiple biclustering algorithms to enhance the stability and accuracy of the results [26].

Recent developments in biclustering have expanded its applicability to a broader range of problems. For instance, biclustering has been used in image processing to segment images into regions with similar textures and colors, and in text mining to identify topics in document collections by clustering words and documents simultaneously [27]. In neuromarketing, these advanced techniques can be leveraged to gain deeper insights into how consumers process visual and textual information in advertisements.

In conclusion, biclustering is a powerful tool for uncovering complex patterns in high-dimensional data. Its ability to simultaneously cluster rows and columns makes it particularly suited for applications in neuromarketing, where it can reveal localized patterns of brain activity related to consumer preferences. By incorporating machine learning and probabilistic methods, modern biclustering techniques offer robust and reliable solutions for analyzing complex datasets.

Biclustering techniques requires one of the following scenarios: either the matrix contains K biclusters, which is typically the number of biclusters we intend to detect, or there is only one bicluster present, as shown in Fig.1.

Figure 1: Bicluster structure

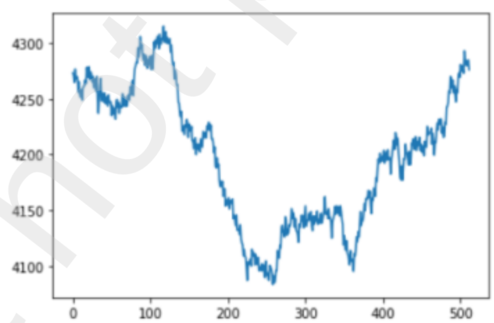


4. Material and Methods

4.1. Applied Method

The research's methodology was divided into four parts (Figure 4). The Emotiv device was used to collect EEG data for the research's initial phase. The Welch transform is applied in the original signal (Figures 2 and 3).

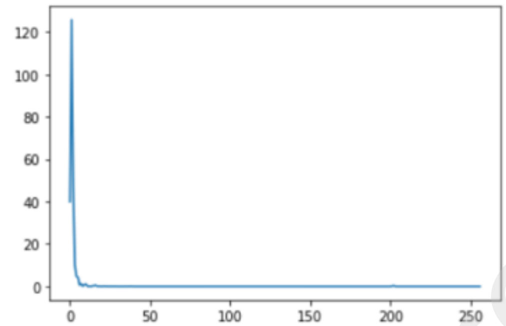
Figure 2: Original signal sample for channel 13



The optimum clustering for training the classification models are chosen using the biclustering technique on the converted signal.

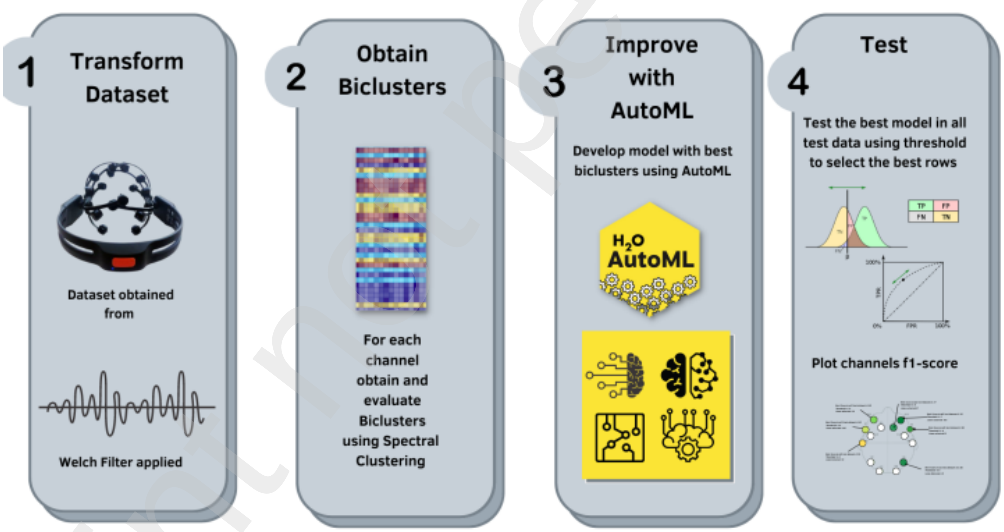
Each bicluster is evaluated with several classifiers and assigned an f1-score for each group. The best biclusters are selected and we use the H2O.ai library to select the best models. Each channel is then analyzed. In contrast to conventional methods, two thresholds are employed for rating: Values over the first threshold show that the customer found the product to be satisfactory. Values lower than the second threshold show that the person

Figure 3: Welch transformed sample for channel 13



did not enjoy the product. The model discards values between the first and second threshold, considered as uncertain.

Figure 4: Applied Method



4.2. DataSet

We used the neuromarketing EEG dataset proposed by [6] for analysis and evaluation. This dataset consists of EEG records of consumers during commercial product evaluation. The authors used the Emotiv Epoch+ device for signal acquisition through 14 channels, following the 10-20 system for

EEG electrodes positioning: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. 25 individuals recorded their EEG signals while looking at consumer goods on a computer screen using all 14 channels. The participants, who are all students at the Indian Institute of Technology in Roorkee, India, range in age from 18 to 38. There are 42 (14 x 3) various product photos in the collection of 14 different products, each of which comes in three different varieties. As a result, a total of 1050 (or 42 x 25) EEG data were captured for each participant. For each photograph in the collection, the subjects' input was gathered in the form of likes and dislikes. EEG signals were simultaneously captured when each image was shown for 4 seconds at a time. The user's preference was recorded after each image was displayed [6].

The internal sampling frequency of the Emotiv Epoch+ device is 2048 Hz, which is downsampled to 128 Hz per channel. The data was transmitted to a computer through Bluetooth using a USB dongle. Participants had an EEG sensor attached to their heads and were asked to view shopping products while their EEG signals were recorded [6]. Figure 5 depicts the evaluated products contained inside this database.

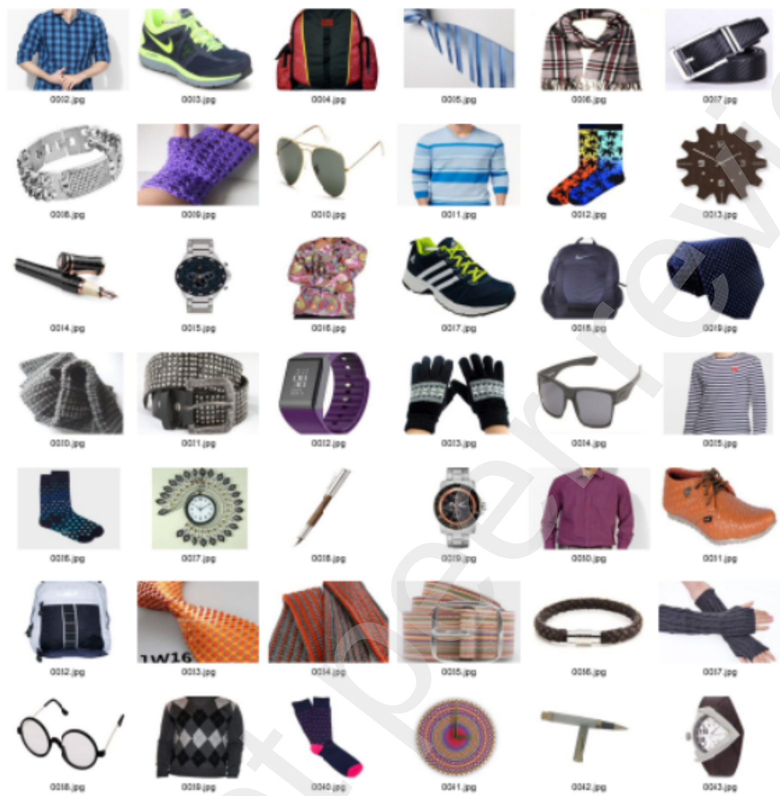
After viewing each product, the participant was asked to evaluate each product with like or dislike categories

4.3. Welch's method

First, Power spectral density (PSD) was applied to the obtained databases. There have been reports of the use of PSD to analyze the EEG signal power in various situations. These include investigating the effects of alcohol on the human brain and validating the neural mass concept. The power spectral density (PSD) illustrates how the power or energy of a sequence is distributed in the frequency domain and is commonly used to measure random signals and noise. The Welch method is one PSD technique that enhances signals at different frequencies. Studies have shown that the Welch approach provides extremely robust EEG signal characteristics, and that these characteristics allow for a clear distinction between classes [28] [29] [30]. Welch's method (also known as the periodogram method) for estimating power spectra divides the time signal into sequential blocks, creates a periodogram for each block, and then averages them.

On a computer, one starts with a series of data values or samples to do spectral analysis [31]: x[0],x[l],...,x[N- 1]. The data sequence's independent variable has a range of 0 to 1. The sample number n is used to index the data values x[n]. This is where the sample value stands in relation to the

Figure 5: Evaluated commercial products used in the studied dataset



sequence's beginning. A consistent pace of data samples acquisition is used. There are T seconds between two consecutive data samples, x[n] and x[n + 1]. There are 1/T samples taken every second. $T_{seq} = N * T$ is the data sequence's length in seconds. The formula $t = t_o + nT$ where is the time when the first data sample was obtained, relates the time of acquisition of a data value to its sample number.

To decompose the data into a collection of weighted sinusoids is the aim of spectral analysis. This decomposition enables one to evaluate the phenomenon under study's frequency content. The phenomenon under examination might be restricted to a specific frequency range. However, it might be dispersed over a wide range of frequencies. There are two main divisions in spectral analysis. A power spectral density or a Fourier transform can be calculated (PSD). Deterministic data is defined as having no random effects or noise. One computes a Fourier transform in this scenario. When random influences mask the intended underlying phenomenon, a PSD is computed. Some type of averaging or smoothing is used to observe the desired phenomenon [31].

Welch's method is also known as the periodogram averaging method and the weighted overlapped segment averaging (WOSA) method. The length of each segment is the parameter M. Segments are sometimes referred to as batches. They refer to M as the batch size. The method use the following steps.

Denote the m th windowed, zero-padded frame from the signal x by [32]

$$x_m(n) \triangleq w(n)x(n + mR), \quad n = 0, 1, \ldots, M - 1, m = 0, 1 \ldots, K - 1 \tag{1}$$

where R is defined as the window hop size, and let K denote the number of available frames. Then the periodogram of the m th block is given by

$$P_{x_m, M(\omega_k)} = \frac{1}{M} |FFT_{N,K}(X_m)|^2 \triangleq \frac{1}{M} |\sum_{n=0}^{N-1} x_m(n) e^{-j2\pi nk/N}|^2 \tag{2}$$

as before, and the Welch estimate of the power spectral density is given by

$$\hat{S}_x^W(\omega_k) \triangleq \frac{1}{K} \sum_{m=0}^{K-1} P_{x_m, M}(\omega_k)$$

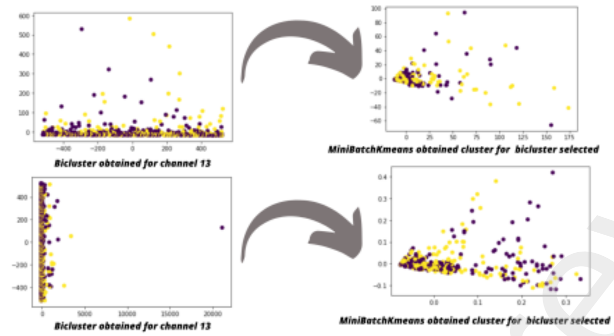
The Welch method from the scipy.signal API is used in thi work.

4.4. Spectral Bi-Clustering

The Spectral Clustering model was used in the second stage to choose the biclusters with the best dimensional separation. Using traditional cluster models like MiniBatchKMeans or MeanShift, the biclusters were further divided.

It has been the subject of extensive machine learning and pattern recognition research to locate suitable clusters. One common method is based on generative models for clustering points in Rn, which is the primary application of this work. In this method, algorithms like EM are employed to learn a mixture density. There are various problems with these methods. First,

Figure 6: Biclusters obtained with MeanShift and MiniBatchKMeans



certain simplification assumptions typically need to be established in order to apply parametric density estimators (e.g., that the density of each cluster is Gaussian). Second, several restarts are necessary to obtain a satisfactory solution using iterative techniques since the log likelihood can have a lot of local minima [33].

Originally proposed for the more frequent one-way clustering of samples (rows of a data matrix), spectral clustering uses a weighted graph where samples are nodes and similarity metrics between samples are used as adjacency weights. It is generally known that employing typical clustering techniques on the eigenvectors of a normalized adjacency matrix or a Laplacian matrix can reveal clusters, as the spectrum of these matrices reflects the clustering or connectedness quality of the underlying graph [34].

The graph partition problem was initially studied using spectral clustering in 1973 [35]. Later, the technique was expanded in [36] and generalized to a variety of applications, including information retrieval [37][38], social networks [39], computational biology [40] 21], medical image analysis, and social networks.

Conventional spectral clustering [2] first calculates a N x N afinity matrix from a dataset of N items, where each entry represents the degree of similarity between two objects as measured by some similarity metrics. The afinity matrix's graph Laplacian is used to conduct the eigen-decomposition, which yields the k eigenvectors and first k eigenvalues. K-means or other discretization approaches can be used to perform the final clustering by embedding the datasets into the low-dimensional space using the k eigenvectors that were obtained [41].

Though finding clusters of arbitrary shapes from complex data using spectral clustering has shown promising benefits, its $O(N)^3$ time complexity and $O(N)^2$ space complexity severely limit its use in large-scale tasks. Some researchers sparsified the afinity matrix using K-nearest neighbors or -neighbors in order to reduce the enormous computational cost. They then used some sparse eigen-solvers to solve the eigen-decomposition problem, which still necessitates computing all of the entries in the original afinity matrix [41].

4.5. AutoML and H2O.ai

Finding and training the best model from the best biclusters obtained is the next step in the suggested process. The idea was to employ the AutoML method.

There have been significant advances made in the creation of user-friendly machine learning software, including scikit-learn, H2O, caret, tidymodels, and mlr, which all offer straightforward, unified interfaces to different machine learning techniques. Although these tools have made it simple for non-experts to train machine learning models, a good deal of knowledge is still necessary to produce cutting-edge outcomes. Tools for automatic machine learning, sometimes known as ”AutoML,” offer a straightforward interface for training numerous models (or a potent single model), making them useful for both beginning and experienced machine learning practitioners [20].

4.6. Neural Architecture Search Problem

Over the past few years, outstanding progress has been made on a variety of tasks, including speech recognition, machine translation, and image and image recognition. The majority of architectures have been created manually by human professionals, which can be a laborious and error-prone process. This has led to a rise in interest in autonomous neural search techniques, or [42].

Automating architecture engineering with Neural Architecture Search (NAS) is the natural next step after automating machine learning. The Neural Architecture Search technique makes an automated effort to find the most suitable architecture and associated parameters for a given issue. On several tasks, such as semantic segmentation, object detection, or picture classification, NAS approaches have already surpassed manually constructed structures. NAS is a branch of automatic machine learning and shares a lot of similarities with hyperparameter optimization and meta-learning. The three criteria we use to classify NAS approaches are research space, research methodology, and performance estimation strategy [42].

Consider F as the search space of a Neural Network and D as the input data divided into D_{train} and D_{val}, with C as the cost function. The objective is to minimize the cost C, with a neural network $f \in F$.

$$f^* = \operatorname{argmin} Cost(f(\theta^*), D_{val}),\ f \in F \tag{3}$$

$$\theta^* = \operatorname{argmin} L(f(\theta), D_{train}),\ \theta \tag{4}$$

Cost is the performance evaluation function for the metric, such as accuracy or mean squared error (MSE), with $\theta*$ as the learning parameter. The search space F encompasses all possible architectures that are morphable from the initial designs.

4.7. Bayesian Optimization

Using the Bayes Theorem is a popular approach to solving an optimization problem. It is a foundational probabilistic model of finding an objective function to achieve a minimum or maximum, according to the desired optimal solution. It consists of testing multiple functions,searching for the final optimal result, which is usually called objective function.

This searching strategy makes this optimization approach suitable for complex problems hard to model, such as machine learning algorithms. It consists in the execution of three main phases. The first is the update phase, consisting of training a general Gaussian model. The second phase is called generation and defines a new model to optimize the solution. Finally, the third phase is the observation, which evaluates the performance of the model created in the previous phase.

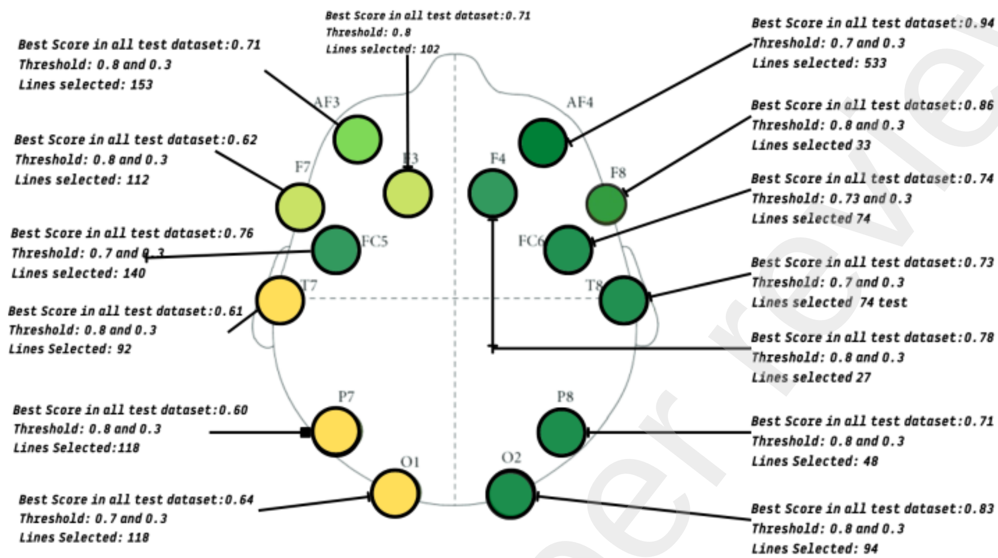
4.8. H2O.ai

H2O is an open source implementation of automated machine learning (AutoML) which aims to provide a fast and scalable platform for intelligent applications. Using automl [43], the API enables complex calculations such as deep learning, nurturing, and bagging ensembles. The instrument includes H2O AutoML, which is a solution to identify possible candidate models to achieve desired solutions or objectives, creating a ranking with the most effective models to solve that specific problem. For that, the AutoML implementation considers different strategies such as gradients, tree/forest-based techniques, and deep learning models, considering the fact that diverse models improve the accuracy of the ensemble method [44]. The OpenML AutoML Benchmark, which analyzes the performance of some of the most well-known open-source AutoML strategies to address the problems of multiple datasets, demonstrates the usefulness of the technique [44].

An automatic machine learning technique called H2O AutoML (H2O.ai, 2017) is part of the H2O framework (H2O.ai, 2013). It is easy to use and produces high-quality models that are appropriate for deployment in a corporate context. Regression, binary classification, and multiclass classification models can be trained under supervision using H2O AutoML'svdata sets in a table. The quick scoring capabilities of H2O models are one of their advantages; several of them can produce predictions in sub-millisecond scoring intervals. H2O AutoML provides APIs for many languages, including R, Python, Java, and Scala, so use is straightforward among a varied group of engineers and data scientists.

Random Forest Classifier its the model with the best f1-score obtained using h2o.ai. The Breiman-proposed random forest classifier [45] is made up of numerous individual classification trees (in this work, there were 10 trees), where each tree is a classifier on its own and is given a set weight for its classification output. The mode (the output with the most votes) of all the classification outputs from the trees is selected in order to determine the overall classification result. The procedures for creating a random forest classifier can be summed up as follows: - Each tree is first constructed before the random forest classifier begins learning. Numerous tree-building techniques, including random trees, CART, J48, REPTree, C4.5, and CHAID, were employed in the literature. The next phase is creating the in-bag data set for each tree, which is created by bootstrapping technique sampling the training data set (sampling with replacement). The nodes and leaves of a random tree classifier were built for each tree using a random number of characteristics.

Figure 7: Scores obtained in each of the 14 channels



4.9. Model Analysis

This paper will consider the confusion matrix and the other three performance indicators to evaluate the proposed classification systems. F1-score, Accuracy, Precision, and Recall.

4.9.1. Confusion Matrix

Let MI(x,y) : $R^2 \rightarrow R$ represent a medical image and $O(MI(x,y)){:}R^2 \rightarrow \Omega, \Omega = 0, 1$ a classification of the medical image MI(x,y). The classification criteria are then presented considering the classification goal as G and the system-achieved result as R [46]:

- true positive: $G(x, y) = 1 \wedge R(x, y) = 1$,
- false positive: $G(x, y) = 0 \wedge R(x, y) = 1$,
- true negative: $G(x, y) = 0 \wedge R(x, y) = 0$,
- false negative: $G(x, y) = 1 \wedge R(x, y) = 0$,

4.9.2. Precision

The Precision is a metric to evaluate the system performance from the perception of the number of false positives, compared to the number of true positive cases, and is given by:

$$P = \frac{TP}{TP + FP}$$

4.9.3. Recall

The Recall (R) is a metric to evaluate the system performance from the perception of the number of false negatives (FN), compared to the number of true positive (TP) cases, i.e., it measures the system performance not to miss one TP case and is given by:

$$R = \frac{TP}{TP + FN}$$

4.9.4. F1-score

The $F1$ score is the harmonic mean between Precision and Recall and is given by:

$$F1 = \frac{2TP}{2TP + FP + FN} \tag{5}$$

4.9.5. Accuracy

Accuracy is a popular metric for evaluating a model's performance, considering both positive and negative classes, and considering true and false positives (TP and FP) together with true and false negatives (TN and FN) in the same metric. The equation is presented as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Two threshold values were established after the outcomes of all phases were tested. For product liking, values over the maximum threshold were positive, while below the lowest threshold the value 0 was used (individual did not like the product). The threshold changed depending on how well each sensor performed on each channel.

5. Experiments and Results

The cerebral cortex that covers the anterior portion of the frontal lobe is known as the prefrontal cortex (PFC). This region of the brain has been associated with complex cognitive planning, personality expression, decision-making, and social behavior control. It is assumed that the primary function of this brain region is the coordination of thoughts and actions in accordance with internal goals. In contrast to the right side of the PFC, which is more focused on avoidance and negative emotions, the left side of the PFC is more biased toward approach, positive goals, and emotions.

The hippocampus, dorsolateral prefrontal cortex, and the midbrain are involved in generating preference for a brand when judgments are based on sensory and cultural information, respectively. The ventromedial prefrontal cortex (VMPFC) is involved in generating preference for a brand when judgments are based solely on sensory information (taste in the case of soda) [47].

Each of the 14 channels was subjected to the proposed method. The channels on the right front face that are displayed had the best results. The best results were obtained by Channel AF4, which had a f1-score of 0.94 and chose 533 lines using the range between 0.7 and 0.3 as valid classifications. The results suggest that the Random Forest model performed better on the signals collected in the right frontal area.

Table 1: Results for all channels

AF4 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.95	0.95	0.95	0.94	317
1	0.93	0.92	0.93	0.94	216
F3 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.72	0.93	0.81	0.72	67
1	0.69	0.31	0.43	0.72	35
AF3 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.88	0.54	0.67	0.71	84
1	0.62	0.91	0.74	0.72	69
O1 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.64	0.85	0.73	0.64	66

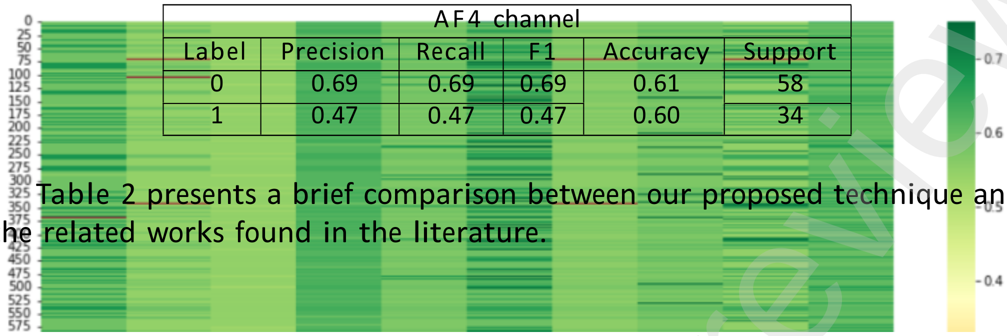
Continued on next page

Table 1: Results for all channels (continued)

AF4 channel					
Label	Precision	Recall	F1	Accuracy	Support
1	0.67	0.38	0.49	0.65	52
T8 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.69	0.78	0.74	0.76	42
1	0.82	0.74	0.78	0.79	42
F8 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.86	1.00	0.93	0.89	25
1	1.00	0.67	0.80	0.91	12
F4 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.91	0.67	0.77	0.78	15
1	0.69	0.92	0.79	0.81	12
P8 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.64	0.82	0.72	0.71	22
1	0.80	0.62	0.70	0.70	26
O2 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.87	0.84	0.85	0.83	56
1	0.78	0.82	0.79	0.83	38
FC5 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.77	0.86	0.81	0.76	83
1	0.75	0.63	0.69	0.76	57
T8 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.74	0.69	0.71	0.73	36
1	0.72	0.76	0.74	0.72	38
FC6 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.72	0.82	0.77	0.74	38
1	0.77	0.67	0.72	0.74	38
F7 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.65	0.55	0.60	0.62	56
1	0.61	0.70	0.65	0.62	56
T7 channel					
Label	Precision	Recall	F1	Accuracy	Support

Continued on next page

Table 1: Results for all channels (continued)



AF4 channel					
Label	Precision	Recall	F1	Accuracy	Support
0	0.69	0.69	0.69	0.61	58
1	0.47	0.47	0.47	0.60	34

Table 2 presents a brief comparison between our proposed technique and the related works found in the literature.

Table 2: Related Works comparison

Work	Apparatus	Dataset	Experiment	Preprocessing Technique	Classification Technique	Best Results
[6]	EEG	Public	e-Commerce Products	Savitzky-Golay Filter	HMM and SVM	Acc: 70%
[7]	EEG	Public	e-Commerce Products	Bandpass filter	SVM, RF and LR	Acc: 51.48%
[8]	EEG	Public	e-Commerce Products	Baseline removal, Bandpass Filter, ICA, PCA, Normalization	SVM	Acc: 66.25%
[9]	EEG	Public	e-Commerce Products	Savitzky-Golay Filter	ANN, LR, DT, KNN, SVM	
[10]	EEG	Public	Mobile brands	ICA	SVM and LDA	Acc: 63%
[11]	EEG	Private	Market and household Product brands	Low-pass filter	NA	NA
[48]	ET, FC, EDA and EEG	Private	Purchase intention	Baseline removal, Bandpass Filter, ICA Normalization	NA	NA
[13]	ET, FC, EDA and EEG	Private	Hair styling products	Bandpass Filter	NA	NA
[49]	EEG	Private	Food commercials	High-pass/Notch Filter, ICA	SVM, LR, Adaboost, KNN	Acc: 68.5%
[50]	EEG	Private	Emotional decisions	NA	NA	NA
This work	**EEG**	**Public**	**e-Commerce Products**	**Savitzky-Golay Filter**	**AutoML**	**Acc: 94%**

e c a n n o t e t h at some recent research works, such as [11, 48, 13] and [50], did not apply classification techniques for predicting like / dislike using EEG signals. Moreover, our results outperform in terms of accuracy those in the literature, possibly predicting consumer preferences from EEG signals.

Figure 8: Bicluster of channel 1 scores

6. Conclusion

Neuromarketing can be understood as the application of neuroscience techniques to understand consumer preferences and behavior. It looks at the brain activity to establish correlation with preferences and purchase intentions. However, these strategies fall short of accurately capturing or explaining the real-time decision-making process of consumers. Traditional marketing strategies take into account a posteriori user input in the form of surveys, product reviews, or commentaries. Additionally, for the validation of product or brand marketing and mass manufacturing, it is essential to comprehend the consumer neurological responses when watching a product advertisement. In this paper, the use of the bi-clustering technique is proposed to minimize noise in EEG signals, followed by automatic machine learning to categorize the brain electric signals. A neuromarketing EEG dataset of consumers' product evaluations for commercial products is used. The 14 channel Emotiv Epoch+ device was used by the authors for data acquisition, following the 10-20 system for the position of the EEG electrodes. 25 participants recorded their EEG signals utilizing all 14 channels while viewing products on a computer screen. The research's methodology was broken up into four sections. For the initial stage of the investigation, the welch transform was applied to acquire EEG data using the Emotiv device. Using the biclustering approach on the transformed signal, the best biclusterings are selected for training the classification models. A different classifier is used to evaluate each biclustering, and an f1 score is given for each group. The best biclustering is chosen and the best models are chosen using the H2O.ai library. Then, each channel is examined. Two thresholds are used for rating as opposed to typical methods: Values above the first threshold indicate that the buyer was satisfied with the product. Lower values of the second threshold indicate that the consumer did not like the product. Because it considers them to be uncertain, the model ignores values between the first and second criteria. By achieving a f1-score value of 0.95, the proposed approach outperforms the present state of the art.

Acknowledgment

This work has been supported by the CNPq via grant numbers 305517/2022-8 and 313599/2019-0 and Coordination for the Improvement of Higher Education Personnel (CAPES)-Finance Code 001.

References

[1] P. M. Oliveira, J. Guerreiro, P. Rita, Neuroscience research in consumer behavior: A review and future research agenda, International Journal of Consumer Studies (2022).

[2] M. M. Niedziela, K. Ambroze, The future of consumer neuroscience in food research, Food Quality and Preference 92 (2021) 104124.

[3] F. S. Rawnaque, K. M. Rahman, S. F. Anwar, R. Vaidyanathan, T. Chau, F. Sarker, K. A. A. Mamun, Technological advancements and opportunities in neuromarketing: a systematic review, Brain Informatics 7 (1) (2020) 1–19.

[4] V. Khurana, M. Gahalawat, P. Kumar, P. P. Roy, D. P. Dogra, E. Scheme, M. Soleymani, A survey on neuromarketing using eeg signals, IEEE Transactions on Cognitive and Developmental Systems 13 (4) (2021) 732–749.

[5] G. M. Wojcik, P. Wierzgala, A. Gajos, Evaluation of emotiv eeg neuroheadset, Bio-Algorithms and Med-Systems 11 (4) (2015) 211–215.

[6] M. Yadava, P. Kumar, R. Saini, P. Roy, D. Dogra, Analysis of eeg signals and its application to neuromarketing, Multimedia Tools and Applications 76 (18) (2017) 19087–19111.

[7] M. Alimardani, M. Kaba, Deep learning for neuromarketing; classification of user preference using eeg signals, in: 12th Augmented Human International Conference, 2021, pp. 1–7.

[8] N. Alnuman, S. Al-Nasser, O. Yasin, Classification of products preference from eeg signals using svm classifier, in: 2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE), IEEE, 2020, pp. 174–179.

[9] A. Ullah, G. Baloch, A. Ahmed, A. Buriro, A. Junaid, B. Ahmed, S. Akhtar, Neuromarketing solutions based on eeg signal analysis using machine learning, International Journal of Advanced Computer Science and Applications 13 (1) (2022).

[10] P. Golnar-Nik, S. Farashi, M.-S. Safari, The application of eeg power for the prediction and interpretation of consumer decision-making: A neuromarketing study, Physiology and Behavior 207 (2019) 90–98.

[11] Q. Ma, M. Wang, Q. Da, The effects of brand familiarity and product category in brand extension: An erp study, Neuroscience Research 169 (2021) 48–56.

[12] J. Ausin-Azofra, E. Bigne, C. Ruiz, J. Marın-Morales, J. Guixeres, M. Alcañiz, Do you see what i see? effectiveness of 360-degree vs. 2d video ads using a neuroscience approach, Frontiers in Psychology 12 (2021) 612717.

[13] U. Garczarek-Bak, A. Szymkowiak, P. Gaczek, A. Disterheft, A comparative analysis of neuromarketing methods for brand purchasing predictions among young adults, Journal of Brand Management 28 (2) (2021) 171–185.

[14] A. Hakim, S. Klorfeld, T. Sela, D. Friedman, M. Shabat-Simon, D. Levy, Machines learn neuromarketing: Improving preference prediction from self-reports using multiple eeg measures and machine learning, International Journal of Research in Marketing 38 (3) (2021) 770–791.

[15] L. Hsu, Y.-J. Chen, Neuromarketing, subliminal advertising, and hotel selection: An eeg study, Australasian Marketing Journal (AMJ) 28 (4) (2020) 200–208.

[16] M. Kahani, R. Amini, S. Heidari, Improved emotion recognition using eeg signals in neuromarketing with deep learning models, Journal of Neuroscience Methods 358 (2022) 109347.

[17] G. Vecchiato, L. Astolfi, F. De Vico Fallani, F. Cincotti, D. Mattia, S. Salinari, F. Babiloni, Neurophysiological tools to investigate consumer's gender differences during the observation of tv commercials, Computational Intelligence and Neuroscience 2011 (2011) 1–9.

[18] Y. Liu, O. Sourina, M. Nguyen, Real-time eeg-based human emotion recognition and visualization, in: 2011 International Conference on Cyberworlds, IEEE, 2011, pp. 262–269.

[19] A. Goshvarpour, A. Goshvarpour, S. Rahati, Application of recurrence quantification analysis on emotional eeg signals during music listening, Brain Informatics 4 (4) (2017) 249–256.

[20] E. LeDell, S. Poirier, H2o automl: Scalable automatic machine learning, in: Proceedings of the AutoML Workshop at ICML, Vol. 2020, 2020.

[21] J. A. Hartigan, Direct clustering of a data matrix, Journal of the American Statistical Association 67 (337) (1972) 123–129.

[22] Y. Cheng, G. M. Church, Biclustering of expression data, Proceedings of the 8th International Conference on Intelligent Systems for Molecular Biology (ISMB) 8 (2000) 93–103.

[23] L. Lazzeroni, A. Owen, Plaid models for gene expression data, Statistica Sinica 12 (2002) 61–86.

[24] S. C. Madeira, A. L. Oliveira, Biclustering algorithms for biological data analysis: a survey, IEEE/ACM Transactions on Computational Biology and Bioinformatics 1 (1) (2004) 24–45.

[25] J. Gu, J. Liu, Bayesian biclustering of gene expression data, BMC Genomics 9 (1) (2008) S4.

[26] W. Ayadi, M. Elloumi, J.-K. Hao, Ensemble biclustering of gene expression data with kpis, IEEE/ACM Transactions on Computational Biology and Bioinformatics 16 (6) (2019) 1967–1978.

[27] A. Banerjee, I. S. Dhillon, J. Ghosh, S. Merugu, D. Modha, Biclustering: Algorithms and applications in data mining and text mining, Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2007) 507–508.

[28] O. K. Cura, M. A. Ozdemir, S. Pehlivan, A. Akan, A dynamic mode decomposition based approach for epileptic eeg classification, in: 2020 28th European Signal Processing Conference (EUSIPCO), IEEE, 2021, pp. 1070–1074.

[29] C. C. W. Fadzal, W. Mansor, L. Khuan, N. Mohamad, Z. Mahmoodin, S. Mohamad, S. Amirin, Welch power spectral density of eeg signal generated from dyslexic children, in: 2014 IEEE REGION 10 SYMPOSIUM, IEEE, 2014, pp. 560–562.

[30] M. A. Naderi, H. Mahdavi-Nasab, Analysis and classification of eeg signals using spectral analysis and recurrent neural networks, in: 2010 17th Iranian Conference of Biomedical Engineering (ICBME), IEEE, 2010, pp. 1–4.

[31] O. M. Solomon Jr, Psd computations using welch's method.[power spectral density (psd)], Tech. rep., Sandia National Labs., Albuquerque, NM (United States) (1991).

[32] P. Welch, The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Transactions on audio and electroacoustics 15 (2) (1967) 70–73.

[33] A. Ng, M. Jordan, Y. Weiss, On spectral clustering: Analysis and an algorithm, Advances in neural information processing systems 14 (2001).

[34] F. Liu, Y. Yang, X. S. Xu, M. Yuan, Mutually exclusive spectral biclustering and its applications in cancer subtyping, bioRxiv (2022).

[35] W. E. Donath, A. J. Hoffman, Lower bounds for the partitioning of graphs, in: Selected Papers Of Alan J Hoffman: With Commentary, World Scientific, 2003, pp. 437–442.

[36] J. Shi, J. Malik, Normalized cuts and image segmentation, IEEE Transactions on pattern analysis and machine intelligence 22 (8) (2000) 888–905.

[37] A.-G. Chifu, F. Hristea, J. Mothe, M. Popescu, Word sense discrimination in information retrieval: A spectral clustering-based approach, Information Processing & Management 51 (2) (2015) 16–31.

[38] B. McFee, D. Ellis, Analyzing song structure with spectral clustering., in: ISMIR, Citeseer, 2014, pp. 405–410.

[39] S. White, P. Smyth, A spectral clustering approach to finding communities in graphs, in: Proceedings of the 2005 SIAM international conference on data mining, SIAM, 2005, pp. 274–285.

[40] W. Pentney, M. Meila, Spectral clustering of biological sequence data, in: AAAI, Vol. 5, 2005, pp. 845–850.

[41] D. Huang, C.-D. Wang, J.-S. Wu, J.-H. Lai, C.-K. Kwoh, Ultra-scalable spectral clustering and ensemble clustering, IEEE Transactions on Knowledge and Data Engineering 32 (6) (2019) 1212–1226.

[42] M. Wistuba, A. Rawat, T. Pedapati, A Survey on Neural Architecture Search 20 (2019) 1–21. arXiv:1905.01392.
URL http://arxiv.org/abs/1905.01392

[43] A. Candel, V. Parmar, E. LeDell, A. Arora, Deep learning with h2o, H2O. ai Inc (2016).

[44] E. Ledell, S. Poirier, H2O AutoML: Scalable Automatic Machine Learning (2020).
URL https://scinet.usda.gov/user/geospatial/#tools-and-software

[45] L. Breiman, Random forests, Machine learning 45 (1) (2001) 5–32.

[46] A. Popovic, M. de la Fuente, M. Engelhardt, K. Radermacher, Statistical validation metric for accuracy assessment in medical image segmentation, International Journal of Computer Assisted Radiology and Surgery 2 (3-4) (2007) 169–181. doi:10.1007/s11548-007-0125-1.

[47] O. T. Zineb, B. Larbi, J. Mohamed, C. Yahya, O. E. M. Hadj, B. Ali, Neuromarketing: Where marketing and neuroscience meet, African journal of business management 5 (5) (2011) 1528–1532.

[48] J. M. Ausin-Azofra, E. Bigne, C. Ruiz, J. Marın-Morales, J. Guixeres, M. Alcañiz, Do you see what i see? effectiveness of 360-degree vs. 2d video ads using a neuroscience approach, Frontiers in Psychology 12 (2021) 612717.

[49] A. Hakim, S. Klorfeld, T. Sela, D. Friedman, M. Shabat-Simon, D. J. Levy, Machines learn neuromarketing: Improving preference prediction from self-reports using multiple eeg measures and machine learning, International Journal of Research in Marketing 38 (3) (2021) 770–791. doi:https://doi.org/10.1016/j.ijresmar.2020.10.005.
URL https://www.sciencedirect.com/science/article/pii/S0167811620301038

[50] L. Hsu, Y.-J. Chen, Neuromarketing, subliminal advertising, and hotel selection: An eeg study, Australasian Marketing Journal (AMJ) 28 (4) (2020) 200–208. doi:https://doi.org/10.1016/j.ausmj.2020.04.009.
URL https://www.sciencedirect.com/science/article/pii/S1441358220300355