ISSN 2007-9737

EEG-Based Classification of Consumer Preferences Using PCA

Mauro Daniel Castillo Pérez[1], Verónica de Jesús Pérez Franco[2], Jesús Jaime Moreno Escobar[1,3,4,*], Hugo Quintana Espinosa [1], Brenda Lorena Flores Hidalgo[1,5], Ana Lilia Coria Páez[6]

[1] Instituto Politécnico Nacional, Escuela Superior de Ingeniería Mecánica y Eléctrica, Zacatenco, Mexico

[2] Instituto Politécnico Nacional, Unidad Profesional Interdisciplinaria de Ingeniería y Ciencias Sociales y Administrativas, Mexico

[3] Instituto Politécnico Nacional, Centro de Investigación en Computación, Mexico

[4] Instituto Politécnico Nacional, Escuela Superior de Cómputo, Mexico

[5] Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Mexico

[6] Instituto Politécnico Nacional, Escuela Superior de Comercio y Admistración, Mexico

jmorenoe@ipn.mx, jmorenoe@cic.ipn.mx

Abstract. This study explores the neural correlates of consumer preferences for functional foods using EEG signals from 83 participants. Using Principal Component Analysis (PCA) for dimensionality reduction and visualization, we identified distinctive brain wave patterns associated with liked and disliked food products. PCA revealed dominant activity in Delta (0.97) and Theta (0.92) waves for preferred foods, indicating strong sensoriemotional interaction, while disliked foods showed reduced Alpha (0.23) and Beta (0.14) activity, reflecting decreased cognitive processing. Statistical validation (70% explained variance using PCA, p ¡ 0.05 in permutation tests) confirmed the robustness. The approach demonstrates how integrating PCA can decode consumer behavior, providing useful insights for neuromarketing and product development, such as optimizing sensory attributes or adapting formulations based on neural profiles. Future work could integrate machine learning for predictive modeling.

Keywords. EEG, PCA, Neuromarketing, Functional Foods, Consumer Preferences.

1 Introduction

Arterial hypertension is a major health issue in Mexico, affecting nearly 40% of adults. This condition, characterized by chronic high blood pres-

sure, increases the risk of cardiovascular diseases, strokes, and kidney problems [14, 11]. It is important to highlight that many people are unaware that they suffer from hypertension, which can complicate treatment and increase the likelihood of severe complications. One of the main causes of hypertension is diet. Diet plays a crucial role in both the development and control of hypertension [37]. Excessive sodium intake, commonly found in salt and many ultra-processed foods, is one of the leading factors contributing to elevated blood pressure. In Mexico, the average daily salt consumption exceeds the World Health Organization's (WHO) recommendation of a maximum of 5 grams per day, which results in approximately 27,700 deaths from cardiovascular causes each year in the country [13, 3]. In addition to sodium, diets high in saturated fats and added sugars typical of ultra processed foods contribute to overweight and obesity, which are conditions that further increase the risk of hypertension. However, a balanced diet that includes fruits, vegetables, whole grains, and legumes provides essential nutrients such as potassium, fiber, and antioxidants, all of which help maintain healthy blood pressure levels.

The Mexican government has implemented various strategies to promote healthier eating habits. One of the most recent measures is the ban on junk food sales in schools, aimed at encouraging healthier eating from an early age. This initiative seeks to reduce the consumption of products high in fat, sugar, and sodium among students. Additionally, campaigns have been launched for the prevention and early detection of non-communicable chronic diseases, with the goal of reducing the prevalence of hypertension, type 2 diabetes, and obesity by 5% to 10% over six years. These campaigns include the monitoring of at least 500,000 individuals to implement effective preventive actions [1, 36].

The objective of this study is to analyze consumer behavior based on their perceptions during decision-making, in order to predict their choice of functional products. This is achieved by using the mathematical tool Principal Component Analysis (PCA) to process data obtained through electroencephalography (EEG). The objective is to identify various brain responses associated with different perceptions and evaluate their relationship with the participants' expressed preferences. To do this, data is collected from a group of individuals who try different functional products, and the brainwave frequencies present during consumption are identified. The results seek to determine whether there are specific brain activity patterns related to consumer decision-making.

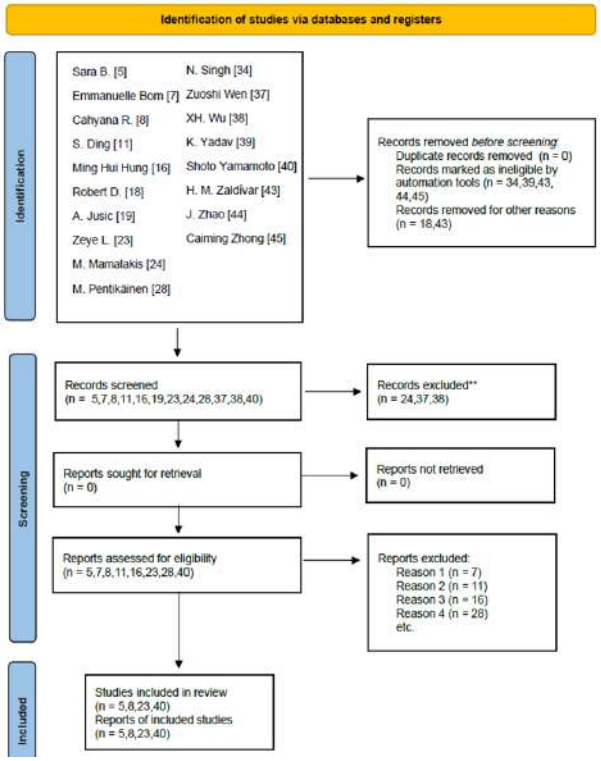

Fig. 1. Identification of Related Works via research databases

To conduct this research, a literature review was conducted to identify studies that use similar tools and methods for data processing, with a focus on efforts to reduce hypertension in Mexico and understand how human brain frequencies respond to different stimuli. The selected articles include recent publications from 2021 to 2025, as well as two studies from the 1990s that began exploring the use of mathematical tools to identify rising levels of hypertension at that time. The articles related to this research are presented in Table 1.

Figure 1 shows the PRISMA flowchart, a selection methodology that includes a systematic review

Table 1. Related Works

Reference	Title	Summary
[39]	Single-cell analysis of peripheral blood from HAPH patients	Identifies increased non-classical and intermediate monocytes, suggesting a role in pathogenesis.
[15]	AI-Based Screening for White-coat and Masked Hypertension	Machine Learning identifies key gut microbiota; SMOTE and SHAP improve model accuracy and interpretability.
[40]	Role of Gut Microbiota in Hypertension	ML reveals microbial taxa linked to hypertension; highlights interpretable models and data balancing.
[7]	Senescent Cells and Pulmonary Hypertension	Analyzes impact of senescent cells in disease using patients and animal models.
[18]	ML Model Based on microRNAs for Essential Hypertension	Uses SVM and miRNAs to create a high-accuracy diagnostic model integrating clinical data.
[10]	LLM Knowledge Distillation for Health Event Prediction	CKLE model improves prediction of heart failure and hypertension by 4.48% using LLM and EHR data.
[46]	Monocyte Pathways in Pulmonary Arterial Hypertension	Shows reduced monocyte infiltration in PAH.
[38]	NAMPT and Macrophage Polarization in PAH	Links macrophages to smooth muscle cell modulation in PAH.
[41]	Facial Image Detection of Hypertension	Combines visible and NIR images using sparse coding; achieves 81% accuracy.
[45]	Dietary Patterns and Hypertension Risk	PCA vs principal balances analysis used to link diets with hypertension risk.
[27]	Treatment Patterns for PAH and CTEPH in Finland	Describes treatment patterns over 12 years in Finnish cohort.
[24]	Transparent AI for Lung Disease in PH	Uses uncertainty estimation and dimensionality reduction in 3D lung models.
[33]	Transcriptomics of Endothelial Cells in PH	RNA-seq reveals PAH-specific signatures from catheter-derived cells.
[17]	Ocular Blood Flow in Essential Hypertension	Measures blood flow changes after Trandolapril treatment.
[44]	Sympathetic Tone in Essential Hypertension	Evaluates sympathetic activity dynamically, showing limitations of baseline measures.

of citations from articles related to the present research to discard those less relevant to the objective and the method used.

Thanks to this process, four articles related to the proposal were identified, each with its own perspective and use of mathematics, whose objective is to contribute to the reduction of hypertension, in addition to using various data reduction methods, especially principal component analysis (PCA). In addition to this method, these works use other techniques such as:

1. t-SNE (t-distributed Stochastic Neighbor Embedding): This is a nonlinear dimensionality reduction technique that preserves the local structure of the data, ideal for two- or three-dimensional visualizations [34].
2. Autoencoder: which is an unsupervised neural network that learns a compressed representation, in other words, an encoding of the input data, which is useful for dimension reduction and noise removal in various studies [5].
3. UMAP (Uniform Manifold Approximation and Projection): is a nonlinear dimensionality reduction method that preserves both the local and global structure of the data, offering high speed and is faster than t-SNE on large data sets. [35].

PCA and t-SNE Implementation for KNN Hypertension Classification Visualization [9], where the authors of the article address the high prevalence of hypertension and the importance of its accurate classification for effective treatment through an approach based on machine learning algorithms and dimensionality reduction techniques, with the aim of predicting the type of hypertension from common characteristics and improve data visualization to increase awareness about the disease, the methodology used is data collection, its classification with K-Nearest Neighbors, in order to apply a dimensionality reduction with Principal Component Analysis and t-Distributed Stochastic Neighbor Embedding, thus having a model whose input is the use of a database of 7,794 patients with characteristics such as age, weight and blood pressure, and then a process with these data, in order to obtain the classification of the types of hypertension at the output and generating understandable visualizations for medical decision making, demonstrating an improvement in the accuracy of the model and in the interpretation of data complex, which may contribute to more effective diagnosis and greater awareness of hypertension.

As another example, there is the article, noncontact remote sensing of abnormal blood pressure using a deep neural network: a novel approach for hypertension screening, [22], where the use of a deep learning model is explored to detect hypertension from infrared thermal images, with the aim of developing a non-invasive detection system that allows rapid detection, given that hypertension often goes unnoticed and can lead to cardiovascular emergencies, it is crucial to have accessible and efficient methods for its early detection, the methodology uses a deep neural network PanycNet to analyze thermal images of different parts of the body, captured by a specialized device, processed with a deep learning model and evaluated with metrics such as precision, recall and AUC; the system receives as input thermal images and clinical data of the participants, processing them to generate results on the presence of hypertension, along with visualizations of the most relevant body regions for classification, thus representing an innovation in hypertension screening with potential applications in public health and safety.

In the article, a comparison of principal component analysis, partial least-squares, and reduced-rank regressions in the identification of dietary patterns associated with hypertension: YaHS-TAMYZ and Shahedieh cohort studies [4], the authors examine the relationship between dietary patterns and hypertension in two parts of a population in Iran, using advanced statistical techniques such as PCA, PLS and RRR to identify and assess these patterns from 32 food groups and their association with hypertension, a significant public health problem; the methodology includes the assessment of dietary intake using a food frequency questionnaire, the use of data reduction techniques to derive dietary patterns, linear regression to assess

the variance explained by these patterns, and independent t-tests and chi-square tests to compare variables between groups.

Input data include participants' demographic information and dietary intake data obtained from the questionnaire, while the process encompasses data collection, identification of dietary patterns and analysis of their relationship with hypertension, thus resulting in significant associations between certain dietary patterns and the risk of hypertension, highlighting diet as a modifiable factor in its prevention, and providing valuable evidence for future research and the design of dietary interventions to reduce the prevalence of hypertension.

Finally, there is the study Hypertension Detection in Facial Image of Visible and Near-Infrared Bands Using Sparse Coding, [41], where the use of a machine learning model based on microRNAs for the diagnosis of essential hypertension is analyzed, with the aim of improving its early detection and clinical management, developing and validating a model that identifies patients with hypertension from microRNA profiles, evaluating accuracy and applicability in different populations; the integration of microRNAs in clinical practice faces challenges in data collection and acceptance by patients and health professionals, the study design included 174 participants, divided into two phases: one of discovery through small RNA sequencing to identify relevant microRNAs, and another of validation through RT-qPCR (Reverse Transcription quantitative Polymerase Chain Reaction, a technique used to convert RNA into complementary DNA and quantitatively measure gene expression in real time); the methodologies used were small RNA sequencing to identify regulated microRNAs in hypertensive patients, their validation by RT-qPCR and the use of a support vector machine model for classification; The input data included blood samples, clinical data, and microRNA levels obtained through the aforementioned techniques; the final result was a highly accurate ML model for classifying patients as hypertensive or non-hypertensive, in addition to identifying specific microRNAs as biomarkers for hypertension, facilitating its early detection and treatment.

This research proposes an innovative approach compared to previous related work, applying advanced mathematical tools such as Principal Component Analysis (PCA) to analyze neural data. Unlike previous studies that also explore this tool to combat hypertension in people, this work focuses specifically on the analysis of brain signals (EEG) obtained while participants observe and taste functional foods. The main objective is to identify the brain frequencies most frequently activated upon contact with these products, allowing us to infer the degree of attraction or preference toward them. This will contribute to the development of more effective strategies to promote the consumption of functional and healthy foods, making them more attractive to consumers. This ties in with the current problem, where excessive consumption of ultra-processed products is associated with an increase in obesity and, consequently, with a higher risk of developing hypertension.

By understanding consumers' brain preferences, products and campaigns can be designed that are more aligned with their neurocognitive responses, representing a significant contribution to both the food industry and public health. The uniqueness of this approach lies in the application of dimensionality reduction (PCA) methods to complex brain data to simplify, visualize, and interpret patterns that would otherwise be difficult to detect.

These methods make it possible to reduce the large volume of variables present in EEG signals, preserving the most relevant information to facilitate the identification of neural responses to food stimuli. Principal Component Analysis (PCA) allows the detection of linear relationships between variables and the generation of new components that concentrate the greatest variability of the original dataset.

This provides a compact and manageable representation that facilitates the interpretation of brain responses. Therefore, the structure of this article is organized as follows: Section 2 describes the mathematical tools used, specifically Principal Component Analysis (PCA), as well as the proposed methodology, detailing the composition of the approach used and the electroencephalography (EEG) signal collection process. Next, Section 4 presents the main findings derived from the application of dimensionality reduction techniques

to EEG signals, in addition to explaining the experimental design and the results obtained during the research. Subsequently, Section 4.4 analyzes and interprets the results, highlighting the patterns identified in brain activity during the evaluation of functional products, with an emphasis on the usefulness of PCA in this analysis. Finally, Section 5 presents the most relevant reflections and general conclusions of the study.

2 Bridging Theory and Practice: EEG Signal Acquisition and Analytical Frameworks for Neuromarketing

This section is structured in three main subsections and addresses brain frequencies from a scientific and analytical perspective. The first subsection, *Mathematical Basis* presents the theoretical foundations, including the role of Principal Component Analysis (PCA) in reducing the dimensionality of data and facilitating the interpretation of brain signals. It also explains what an electroencephalogram (EEG) is and its relevance for recording brain activity.

The second subsection, *Method*, describes the techniques used for signal acquisition and processing, focusing on how the EEG signals were obtained and their importance for analyzing brain activity related to decision-making. The general operating model of the proposed system is also presented.

Together, these subsections provide a comprehensive and rigorous framework for studying the interaction between brain processes and consumer decision-making, an essential aspect for the advancement of neuromarketing.

2.1 Mathematical Basis

2.2 Definition of Principal Component Analysis

Principal Component Analysis (PCA) is a statistical technique that helps simplify large datasets by converting them into a smaller number of uncorrelated variables known as principal components. These components are linear combinations of the original variables and are designed to capture as much variance as possible from the dataset [8, 29].

PCA transforms the original data into a new coordinate system using eigenvectors and eigenvalues of the covariance matrix. The eigenvectors indicate the directions in which the data exhibits the greatest variance, while the eigenvalues quantify the importance of these directions. The general steps in applying PCA are:

1. Standardization: Data is normalized to avoid biases that may arise from differences in scales, since PCA is sensitive to the variances of the initial variables. That is, if there are large differences in the ranges of the original variables, those with greater ranges will dominate over those with smaller ranges, leading to biased results.

2. Covariance Matrix Calculation: This step identifies the correlation between variables. Sometimes variables are highly correlated, thus containing redundant information. To detect these correlations, the covariance matrix is computed. This matrix is symmetric of size $P \times P$ (where P is the number of dimensions) and contains the covariances associated with all possible pairs of the original variables:

$$\mathbf{F} = \begin{bmatrix} (X,X) & (X,Y) & (X,Z) \\ (Y,X) & (Y,Y) & (Y,Z) \\ (Z,X) & (Z,Y) & (Z,Z) \end{bmatrix}. \quad (1)$$

3. Eigenvector and Eigenvalue Calculation: In this step, the directions and the amount of variance contained in the data are determined. These values are the core of the magic behind principal components. The eigenvectors of the covariance matrix represent the axes directions where the greatest variance (the most information) exists, [19, 25], and these are called principal components. The eigenvalues are the coefficients associated with the eigenvectors that indicate the amount of variance present in each principal component.

Principal Component Analysis is an incredibly powerful and versatile statistical technique with a wide variety of applications in research and data analysis. Thanks to its ability to simplify complex data and reveal patterns, PCA has become a

fundamental tool for exploring and analyzing large datasets [28, 12].

2.3 Definition of Covariance

Covariance is a statistical measure that quantifies the degree to which two random variables vary together. It is widely used in probability theory, statistics, and machine learning to understand the relationship between features or signals. If the variables tend to increase or decrease together, the covariance will be positive; if one increases while the other decreases, it will be negative. A covariance close to zero indicates that the variables are linearly independent [23].

Mathematically, the covariance between two variables X and Y with n observations is defined as:

$$\text{Cov}(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}), \qquad (2)$$

where:

— x_i and y_i are the individual observations of X and Y,

— $\bar{x}$ and $\bar{y}$ are the means of X and Y,

— n is the total number of paired observations.

The computation can be expressed in matrix form for a dataset $\mathbf{X}$ of size $n \times m$ (where m is the number of features):

$$\mathbf{\Sigma} = \frac{1}{n-1}(\mathbf{X} - \bar{\mathbf{X}})^T(\mathbf{X} - \bar{\mathbf{X}}). \qquad (3)$$

Here, $\mathbf{\Sigma}$ is the covariance matrix, $\bar{\mathbf{X}}$ is the mean vector replicated for each observation, and each element σ_{ij} in $\mathbf{\Sigma}$ represents the covariance between features i and j.

The covariance matrix has important properties:

1. It is symmetric: $\sigma_{ij} = \sigma_{ji}$.
2. The diagonal elements σ_{ii} are the variances of each variable.
3. Positive, negative, or zero covariances indicate the direction and strength of the linear relationship between variables.

To illustrate, suppose we have EEG signals from two electrodes X and Y. The covariance can reveal whether increases in signal amplitude from electrode X tend to coincide with increases or decreases in electrode Y. This measure is essential in identifying correlated brain regions or redundant features before applying dimensionality reduction techniques.

2.3.1 Applications

In machine learning, covariance is fundamental in:

1. Feature selection — identifying redundant or highly correlated features.
2. Principal Component Analysis (PCA) — which directly uses the covariance matrix to find directions of maximum variance.
3. Time-series analysis — assessing the synchrony between signals.

Covariance thus serves as a cornerstone for understanding relationships within data and guiding further dimensionality reduction or preprocessing steps.

2.4 Definition of Electroencephalogram

The electroencephalogram (EEG) is a technique that allows us to record the electrical activity of the brain. This activity is generated when the dendrites of neurons in the cerebral cortex are stimulated through synapses; this stimulation causes electrical impulses to be emitted in the dendrites, creating magnetic and electric fields that can be detected by EEG systems. This enables the analysis of brain function both under normal conditions and during various tests [21, 2].

However, the signal weakens due to the different layers that make up the human head, such as the brain, skull, and scalp, which can also introduce both internal and external interference. Therefore, only a considerable number of active neurons can generate a potential strong enough to be recorded by surface electrodes. Nonetheless, brain electrical signals can be detected in various layers of the human brain, such as the scalp, the base of the skull with the brain exposed, or in deeper areas of

the brain. For this, electrodes can be superficial, basal, or even surgically implanted, with superficial electrodes being the most common in EEG practice [42].

In the field of brain-computer interfaces (BCI), surface electrodes are key for recording brain activity. Several acquisition methods exist, such as the Illinois system, Montreal Aird, and Lennox, among others. The most widely used standard in research is the International 10-20 System. This positioning system is based on specific anatomical landmarks, such as the inion, nasion, and ear lobes, ensuring uniform electrode placement regardless of head size. Its name comes from the distribution of electrodes at intervals of 10% or 20% of the reference anatomical distances [26].

2.5 Acquisition of EEG Signal

An EEG signal represents the electrical activity generated when the dendrites of pyramidal neurons in the cerebral cortex are synaptically stimulated. This neuronal activation produces electrical impulses that, in turn, generate electric and magnetic fields detectable by electroencephalography (EEG) systems; an example of this electrical impulse can be seen in Figure 2. The human skull is distinguished by three main layers: the brain, the cranium, and the scalp, which attenuate the signal and can introduce internal and external interference. For this reason, only a large number of simultaneously active neurons can generate a potential intense enough to be recorded by surface electrodes [32, 43].

The electrical signals produced by the brain can be captured on the scalp, at the base of the skull with the brain exposed, or in deeper brain areas. The electrodes used for detection can be superficial, basal or surgically implanted, with superficial electrodes being the most commonly used in EEG studies [30, 31].

In brain-computer interface (BCI) systems, surface electrodes placed on the scalp are commonly used to record brain electrical signals. There are several acquisition methods, also known as arrays or positioning systems, such as those developed in Illinois, Montreal Aird, and Lennox, among others. However, the most widely used method in research—including this study—is the International 10-20 Positioning System [6].

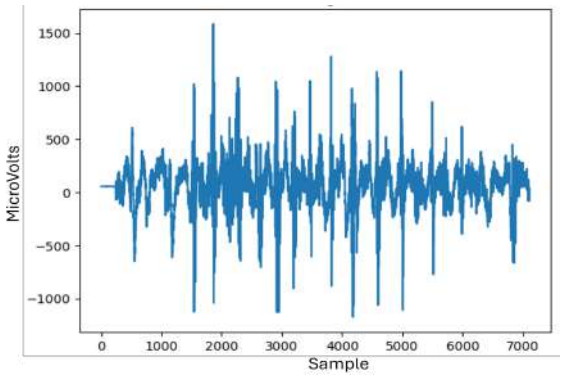

Fig. 2. Example of a cerebral electrical impulse

The International 10-20 System is a standardized protocol that relies on specific anatomical landmarks such as the nasion, inion, and earlobes to ensure consistent electrode placement regardless of head size. Its name comes from the fact that electrodes are placed 10% to 20% away from these anatomical landmarks.

This study uses the ThinkGear TGAM1 EEG device, which detects neural signals and sends them to the ThinkGear chip for processing, generating a stream of useful data and digitally removing potential interference. The raw brain signals are amplified and analyzed to provide relevant information to the system. Acting as a link between the brain and the computer, the NeuroSky ThinkGear module, shown in Figure 3, was selected for being a non-invasive technology with a reliability level of 98%. The electrode is placed at location F_{p1} [20, 16].

3 Methodology

3.1 General Model

The component analysis system, Figure 4 uses mathematical dimensionality reduction tools, such as PCA and t-SNE, to extract the principal components of a person's EEG signal. To do this, the EEG signal is input and processed using these statistical techniques, which transform and analyze

Fig. 3. Sensor ThinkGear ASIC Module v1.0 (TGAM1)

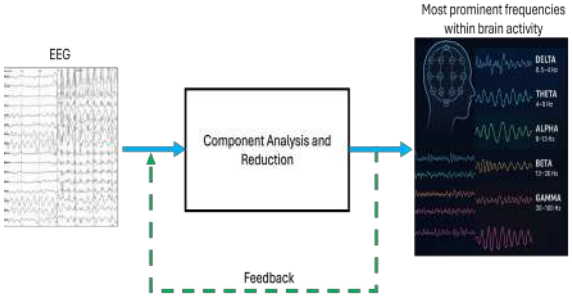

Fig. 4. General model of the system

the data to identify the predominant frequencies of brain activity. In addition, the model has a feedback mechanism that allows the mathematical analysis to be dynamically adjusted based on the results obtained, thereby improving its accuracy.

3.2 The Internal Model of the System

Internally, the model performs several processing steps, beginning with preprocessing the EEG signal, which involves filtering and cleaning the data to reduce noise and corrupted signal components. Then, upon moving to the next block, the signal is transformed into the frequency domain, leading to the next block where the signals are classified. The resulting data are then processed using two dimensionality reduction techniques: Principal Component Analysis (PCA) and t-SNE. These techniques operate in parallel to extract meaningful patterns, which are then combined to form the principal frequency components in the final block. The output is thus the identification of the predominant frequencies of brain activity, Figure 5.

3.2.1 Signal Preprocessing

The first block in the processing pipeline is the Signal Preprocessing stage. Its main purpose is to clean and prepare the raw EEG signal for further analysis. EEG signals are typically contaminated with various types of noise and artifacts, such as muscle activity, eye movements (like blinking), and external electrical interference, especially from the power line (50 or 60 Hz).

Because of this, preprocessing is a crucial step to ensure the reliability of the data before applying any transformations or analyses.

The preprocessing process usually begins with a bandpass filtering, which preserves only the frequency range of interest, generally between 0.5 Hz and 100 Hz. This can be mathematically represented as:

$$x_{\text{filtered}}(t) = \text{BandpassFilter}(x(t), f_{\text{low}}, f_{\text{high}}), \quad (4)$$

where $f_{\text{low}} = 0.5$ Hz and $f_{\text{high}} = 100$ Hz.

Finally, a normalization or standardization process is applied, adjusting each signal segment to have zero mean and unit variance. This makes it easier to compare data across different samples or subjects. The formula used is:

$$x_{\text{norm}} = \frac{x - \mu}{\sigma}. \quad (5)$$

As a result of these procedures, we obtain clean, normalized, and well-structured EEG signals, ready to be transformed.

The next step in the pipeline is the Domain Transformation block, which is responsible for converting signals from the time domain into the frequency domain. This allows for a deeper analysis of brain activity.

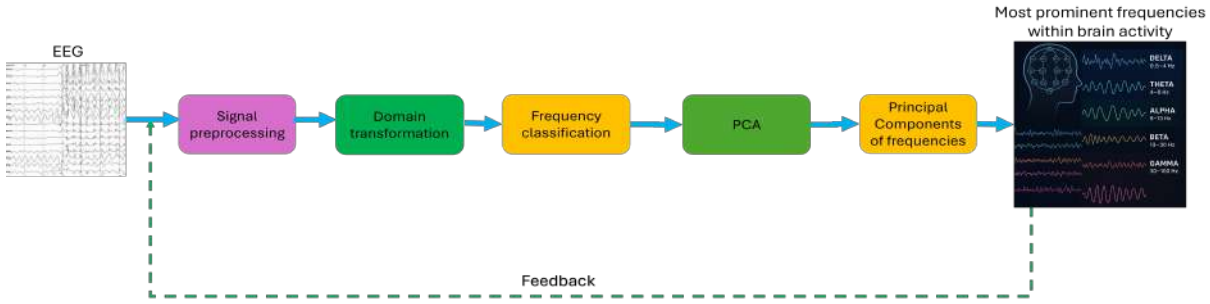

Fig. 5. Internal model of the system

3.2.2 Domain Transformation

The second block in the EEG processing pipeline is known as Domain Transformation. Its main objective is to convert the EEG signal from the time domain where it was originally recorded—into the frequency domain. This transformation is crucial for analyzing the brain's electrical activity in terms of its constituent frequency bands: Delta, Theta, Alpha, Beta, and Gamma, which are associated with various cognitive and physiological states such as attention, relaxation, and mental workload.

To achieve this transformation, mathematical tools like the Fast Fourier Transform (FFT) or the Wavelet Transform are typically used. Each technique has its advantages depending on the type of analysis required.

The FFT provides a global view of the frequency content of the signal, under the assumption that it is stationary. It is defined as:

$$X(f) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j2\pi fn/N}, \tag{6}$$

where $x(n)$ represents the EEG signal in the time domain, and $X(f)$ is its representation in the frequency domain.

However, EEG signals are often non-stationary, meaning their properties vary over time. In these cases, the Wavelet Transform either Continuous (CWT) or Discrete (DWT)—is more suitable, as it provides both temporal and frequency resolution. The continuous wavelet transform is expressed as:

$$W_x(a,b) = \int_{-\infty}^{\infty} x(t) \cdot \psi^* \left(\frac{t-b}{a} \right) dt. \tag{7}$$

Here, ψ is the mother wavelet function, which analyzes the signal at different scales a (related to frequency) and translations b (related to time).

As a result, we obtain a frequency-based representation of the signal, which highlights how the energy is distributed across the different brainwave bands. This spectral information is then passed on to the next processing block, usually involving classification or interpretation.

3.2.3 Frequency Classification

The third essential stage in the EEG signal processing pipeline is Frequency Classification. After the signal has been transformed into the frequency domain, this block aims to categorize the signal's spectral content into well-established EEG frequency bands. These bands reflect different brain states and cognitive activities:

1. Delta (δ): 0.5–4 Hz – associated with deep sleep and unconscious states.
2. Theta (θ): 4–8 Hz – linked to relaxation, meditation, and light sleep.
3. Alpha (α): 8–13 Hz – related to calm wakefulness and closed-eye rest.
4. Beta (β): 13–30 Hz – connected to active thinking, attention, and problem-solving.
5. Gamma (γ): 30–100 Hz – associated with higher cognitive functions like perception and memory integration.

To classify the frequency content into these bands, the spectral power of the signal is first estimated. One common method to obtain this is through the Power Spectral Density (PSD), which provides the distribution of signal power across frequencies. It is defined as:

$$P(f) = \frac{|X(f)|^2}{N}. \tag{8}$$

Here, $X(f)$ is the Fourier transform of the time-domain EEG signal, and N is the number of frequency bins or samples.

Once the PSD is computed, the next step involves integrating the power within each frequency band. This quantifies the total energy associated with each brainwave range. The power within a specific band is calculated using:

$$P_{\text{band}} = \int_{f_1}^{f_2} P(f)\, df, \tag{9}$$

where f_1 and f_2 are the lower and upper bounds of the respective EEG band.

The final outcome of this classification process is a feature vector for each subject or signal segment. This vector consists of either the absolute or relative power values in the five EEG frequency bands:

$$\vec{v}_{\text{freq}} = [P_\delta, P_\theta, P_\alpha, P_\beta, P_\gamma]. \tag{10}$$

This vector is then passed to the next processing block, which performs dimensionality reduction using techniques such as Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE).

4 Results

4.1 Experimental Setup

For the execution of the proposed system, a series of hardware and software requirements were defined, necessary for the execution and acceptable performance of the system during the processing and analysis of EEG signals. First, a Windows 10 operating system is required, which provides a stable environment compatible with the tools used. Regarding the processor, an Intel Core i5 processor is recommended as a minimum, offering the computing power necessary to efficiently execute analysis and machine learning processes.

The minimum RAM to be used is 8 GB, sufficient to handle the volumes of data generated and process multiple tasks simultaneously without affecting system performance. For the acquisition of brain signals, the NeuroSky ThinkGear device was used, a portable, low-cost, and easy-to-implement EEG sensor, ideal for experimental studies that require efficient and simple capture of brain electrical activity. Data processing and the execution of the mathematical tools necessary for analysis were performed using the Google Colab platform. This platform was chosen for its ability to offer cloud computing resources, such as graphics processing units (GPUs) and preconfigured development environments in Python, which significantly facilitates and accelerates model development and execution.

Finally, the system was experimentally validated with the participation of 81 test subjects, all university students between the ages of 18 and 25. Each participant consented to the collection and use of their EEG signals, thus contributing to the development of a way to identify which brain frequencies are most prominent when making a decision.

4.2 Method under Study

The proposed system works with two output classes: like and dislike, which represent the user's preference for a healthy food. Brain frequency classification is achieved through data reduction and analysis using Principal Component Analysis (PCA), which allows for the extraction of relevant patterns from the processed EEG signals. This dimensionality reduction tool helps identify the principal components associated with the participant's brain activity, thus facilitating the grouping of responses into like or dislike categories.

4.3 PCA

4.3.1 Like

EEG Frequency Band Correlation Matrix

Figure 6 shows an analysis of the correlation matrix of EEG frequency bands, revealing significant

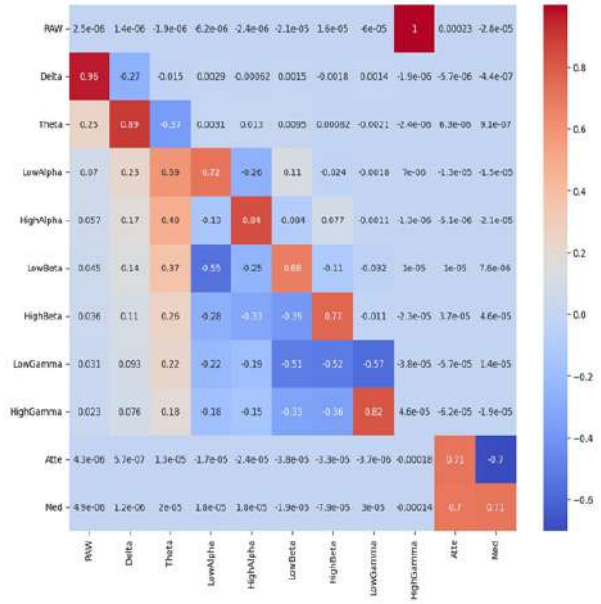

Fig. 6. Inference Matrix "Like"

associations between different frequencies. The Delta band shows a high correlation (0.97) with the raw EEG signal (RAW), indicating a strong influence on the recorded brain activity. Similarly, Theta presents a high correlation with Delta (0.92), indicating that both share similar patterns. On the other hand, the High Alpha band shows a considerable correlation with Low Alpha (0.75), implying a close relationship between the two during brain activity. Likewise, a positive correlation is observed between High Gamma and Attention (0.69), indicating an association with cognitive focus. Additionally, Attention and Meditation are positively correlated (0.72), suggesting a relaxed mental state but with a high level of concentration.

On the other hand, negative correlations are shown, such as Low Gamma (-0.6) and High Gamma (-0.44), which show inverse relationships with other bands, indicating neuronal behavior contrary to lower frequencies. Similarly, the negative correlation between Attention and Meditation (-0.69) may imply that these mental states operate under opposite dynamics in the EEG signal. During meditation, there is a decrease in Gamma activity, reinforcing the idea that higher Gamma activation is associated with less relaxation.

From these correlations, it is concluded that Delta and Theta bands dominate the EEG signal during the consumption of pleasant functional foods, reflecting sensory and emotional processing. Alpha and Beta bands, for their part, show interactions related to the cognitive evaluation of food, including its flavor, texture, and acceptance. Gamma waves and Attention are positively correlated, suggesting greater cognitive involvement during food enjoyment. The inverse relationship between Attention and Meditation highlights the existence of different neural mechanisms for active focus and relaxed state.

Role of EEG Frequency Bands in the Eating Experience

Each frequency band is associated with specific cognitive and emotional processes, and in the context of food consumption, these associations help to understand at a neural level why participants enjoyed the product. Starting with the Delta band (0.5 – 4 Hz), it is highly correlated with Theta (0.92) and with the raw signal (0.97), and is related to unconscious processing, sensory integration, and reward mechanisms. Theta (4 – 8 Hz), in addition to its correlation with Delta, presents a negative relationship with Alpha (-0.32) and is linked to memory and emotional processing, indicating an emotional connection with the food, evoking memories or creating positive associations.

The Alpha band (8 – 12 Hz) is divided into Low Alpha (0.71) and High Alpha (0.75). Low Alpha is associated with relaxation, while High Alpha reflects greater cognitive engagement, suggesting a pleasurable but also mentally active experience. The Beta band (12 – 30 Hz), with subcomponents such as Low Beta (0.65) and High Beta (0.87), is related to decision-making and cognitive processing, indicating an active evaluation of the product in terms of taste, texture, and quality. Finally, the Gamma band (30 – 100 Hz), in its Low (-0.6) and High (0.74) variants, is linked to higher cognitive functions and sensory perception. High Gamma is also correlated with Attention (0.69), reflecting focus and perceptual intensity during the eating experience.

Cognitive and Emotional Involvement: Attention and Meditation

The analysis shows that Attention and Meditation are positively correlated (0.72), suggesting that participants experienced a state of relaxation accompanied by high focus. However, there is a negative correlation between Meditation and High Gamma (-0.69), implying that higher cognitive activation is associated with lower relaxation. This allows us to interpret that, while enjoying the food, participants combined active cognitive engagement (which could be seen as the evaluation of the food) with a pleasurable emotional response (related to a memory or sensation).

Statistical Validation of PCA in EEG

The Principal Component Analysis (PCA) tool was used to reduce the dimensionality of EEG data and extract significant patterns during the consumption of preferred functional foods. The initial correlation matrix, Figure x, showed strong associations, such as Delta-Theta (0.92) and Low Alpha-High Alpha (0.75), validating the internal structure of the data and justifying the application of PCA. The first principal components explain between 80% and 90% of the total variance, indicating that this technique adequately represents the EEG signal without substantial information loss. Likewise, the factor loadings analysis revealed high loadings in Theta and Beta, key elements in the consumption experience due to their relationship with memory, decision-making, and emotions. Additionally, a negative loading in Gamma during meditation was confirmed, supporting the association between higher Gamma activation and lower relaxation.

Practical Applications

This analysis confirms that EEG signals are capable of capturing neural responses linked to food preferences. The results suggest that food enjoyment involves increased cognitive and emotional engagement. These conclusions are relevant and applicable in various fields such as neuromarketing, sensory analysis, and the development of new food products, offering an objective tool to evaluate consumer experience from a neurophysiological perspective.

4.3.2 Dislike

Correlation Matrix Based on PCA Analysis for Disliked Functional Foods

The figure 7 shows the correlation matrix based on Principal Component Analysis (PCA), indicating the correlations between EEG frequency bands recorded in 83 participants who expressed dislike towards a functional or healthy food. The analysis reveals significant relationships between brain activity, cognitive processing, and emotional responses, providing relevant information for consumer neuroscience and the development of food products.

Firstly, Delta and Theta waves show a strong correlation (0.89), suggesting co-activation during moments of deep cognitive processing and memory retrieval. This pattern suggests that participants may have compared the taste of the food with previous negative experiences, which is consistent with studies linking theta activity to emotional evaluation and episodic memory. On the other hand, both High and Low Alpha waves are strongly correlated with each other (0.72 and 0.84), suggesting a synchronized decrease in sensory perception and attention. This dynamic may indicate a state of cognitive disconnection or suppression in response to the rejection of the product.

Beta waves, both Low and High, also show a positive correlation (0.77), indicating active information processing. This pattern suggests that participants were critically evaluating the product, possibly judging its taste, texture, or quality, reinforcing their negative perception. Regarding Gamma waves, a high inter-correlation (0.82) is observed between Low and High Gamma, indicating an increase in emotional response and multisensory integration—a phenomenon commonly associated with rejection responses.

In terms of mental state, an inverse correlation (-0.7) was identified between attention and meditation levels. This indicates that greater focus on the negative sensory stimulus results in lower relaxation and acceptance. Both indicators, attention and meditation, also show correlations with the

ISSN 2007-9737

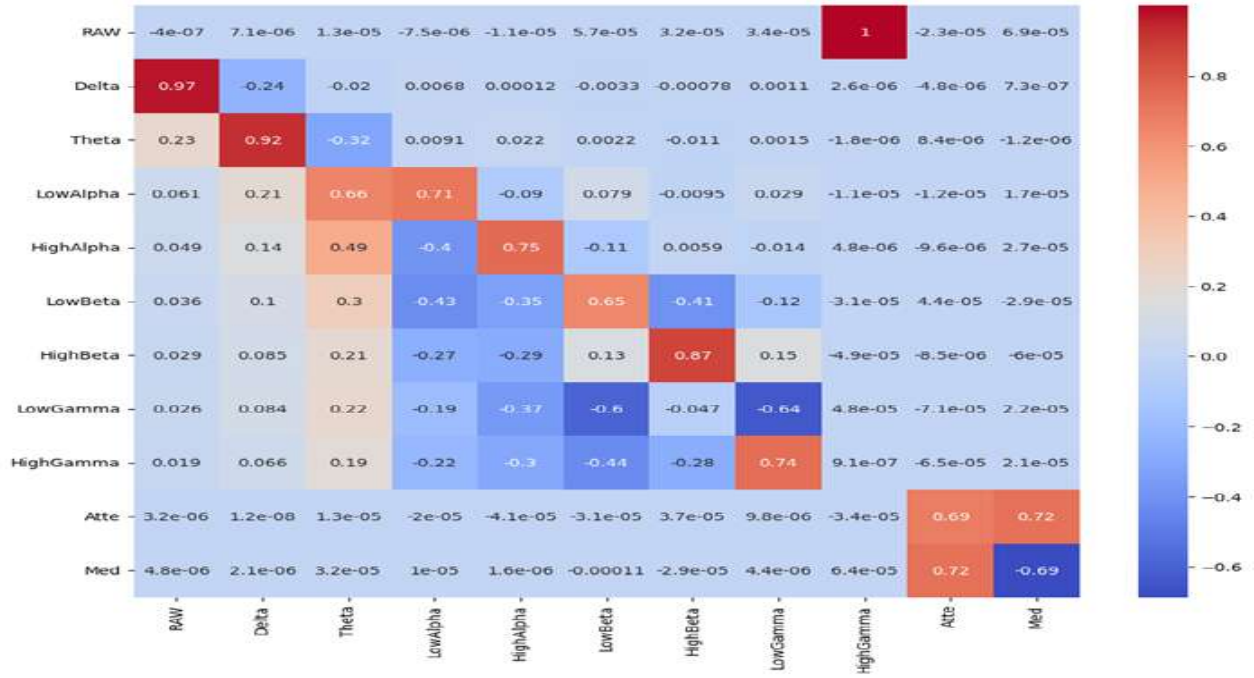

Fig. 7. Inference Matrix "Dislike"

high gamma band, reinforcing the idea that participants experienced an active but negative cognitive-emotional state.

The interpretation of the analysis suggests that when participants rejected the functional product, specific brain mechanisms were activated: the synchronization between Delta and Theta indicated memory and emotional evaluation; the increase in Beta activity reflected cognitive tension and critical judgment; and the intensification of Gamma denoted sensory discomfort. The inverse relationship between attention and meditation describes an active yet unpleasant mental state. These results demonstrate the value of EEG in capturing subconscious emotional reactions to food products, which is highly relevant for neuromarketing strategies.

4.4 Discussion

The joint analysis of EEG brain responses from participants who expressed liking and disliking toward functional foods reveals significantly distinct neural patterns, reflecting divergent emotional and cognitive processes. Through Principal Component Analysis (PCA), correlations were identified among frequency bands that allow inference of underlying mechanisms of acceptance or rejection of these types of products.

In cases where participants reported liking, notable synchronization was observed between high and low Alpha waves, as well as between Theta and Alpha waves. This suggests a state of relaxation, positive emotional processing, and sensory openness. This pattern aligns with favorable affective processing, in which the gustatory experience is associated with pleasant sensations or positive memories. The positive correlation between attention and meditation, along with their association with gamma activity, indicates a state of full but pleasant mental engagement. This could be related to a harmonious sensory integration and a favorable cognitive evaluation of the product.

In contrast, participants who reported disliking exhibited an opposite pattern of brain activity. The high correlation between Delta and Theta bands suggests deeper cognitive processing, linked to

the evocation of possibly negative memories or critical emotional evaluation. The increase in Beta and Gamma waves indicates cognitive tension, critical evaluation, and heightened emotional arousal—typical features of an aversive or rejection response. Furthermore, the inverse relationship between attention and meditation reflects an active but uncomfortable mental state, where attentional focus is directed toward an unpleasant sensory stimulus, thereby reducing subjective well-being.

These differences not only reflect distinct neurological processes but also different cognitive and emotional strategies in response to the same type of stimulus (functional food). Liking appears to activate a reward and relaxation circuit, promoting sensory acceptance, whereas disliking activates mechanisms of negative evaluation, alertness, and emotional disconnection—likely as a form of biological protection or rejection. Both patterns offer valuable insights for consumer neuroscience, especially in the context of healthy food products that often generate mixed reactions due to their taste, texture, or association with dietary restrictions.

From an applied perspective, these findings suggest that the acceptance of functional products depends not only on their organoleptic properties but also on their impact on the neural networks involved in gustatory and emotional experience. The use of EEG and PCA has proven to be a robust tool for capturing these subconscious nuances. Statistical validation using scree plots, explained variance, Hotelling's T^2 test, and permutation tests supports the reliability of the results. Moreover, industrial case studies (Nestlé, PepsiCo, Coca-Cola) show that these findings can be translated into concrete strategies for product development, emotional marketing, and consumer personalization.

To reinforce these results, a permutation test was conducted to statistically validate the variance explained by the PCA. The results showed that the original model variance was 1.000000×10^{0}, while the average variance obtained through permutations was 9.090909×10^{-2}. This contrast yielded a p-value of 0.000000, indicating that the patterns detected in the original data are not due to chance. Thus, the null hypothesis can be rejected, and it can be concluded that the identified principal components represent real and significant structures in the EEG data.

This type of validation is essential in contexts such as this, as it ensures that the observed brain differences between liking and disliking conditions are responses to specific stimuli and not random noise. Factors such as the chosen significance level, the number of permutations (1000), EEG data preprocessing, and correction for multiple comparisons were also carefully considered to ensure the reliability of the analysis.

Additionally, the factor loadings of the principal components were analyzed, which reveal how each original variable (EEG frequencies, attention, meditation, etc.) relates to the extracted components. High loadings (positive or negative) indicate a strong influence of the variable on the construction of the component. For instance, a high loading of the 'RAW' signal in PC1 indicates that this signal plays a key role in the structure of that component. These relationships can be visualized using a heatmap, facilitating the identification of key neural patterns.

To complement this, Hotelling's T^2 statistic—a robust multivariate test—confirmed the existence of significant differences between the groups. With a T^2 value of 30080.04, an F-value of 2734.54, and a p-value of 1.11×10^{-16}, the results indicate a clear separation between EEG responses under liking and disliking conditions, further strengthening the model's validity and its practical implications.

K-Fold cross-validation was implemented to ensure the model's stability, finding that the optimal value of K was 2. This configuration allowed for evaluation of the model's generalization to new datasets, with no evidence of overfitting.

Regarding the proportion of variance explained by each component, the following values were obtained:

$$\begin{bmatrix} 0.8536 & 0.1187 & 0.0161 & 0.0054 \\ 0.0029 & 0.0015 & 0.0010 & 0.0005 \\ 1.3 \times 10^{-7} & 2.3 \times 10^{-9} & 5.0 \times 10^{-10} & \end{bmatrix}. \quad (11)$$

The first component explains over 85% of the total variance, and together with the second, they

account for approximately 97%. This notable dimensionality reduction capability without significant information loss enables a clear representation of the differentiated brain patterns between positive and negative emotional states toward the evaluated foods.

Finally, several future research directions are proposed. One is to analyze how different sensory stimuli such as bitterness, sweetness, or even texture relate to specific EEG responses that help identify rejection elements. Further exploration of the implications of Delta and Theta activity is suggested, distinguishing whether they represent relaxation or emotional dissatisfaction. Additionally, it would be useful to examine how individual factors such as age, gender, dietary habits, or lifestyle influence neural responses. Lastly, long-term effects should be studied to observe whether repeated exposure to the product leads to habituation or even progressive acceptance.

5 Conclusions

This study analyzed consumer behavior based on their taste perception, using electroencephalography (EEG) data and techniques such as Principal Component Analysis (PCA) to identify brain patterns associated with different flavors and their relationship to stated preferences. The results show that EEG is an effective tool for detecting specific brain responses linked to the decision-making process, allowing for a better understanding of preferences and the design of more appealing functional products. The use of mathematical tools to synthesize and manage the recorded variables facilitated an accurate and efficient analysis, confirming the viability of this approach for optimizing the development of products tailored to consumer tastes.

As future work, we propose the inclusion of advanced mathematical methods such as t-SNE, autoencoders, and UMAP, which could optimize the dimensionality reduction process, improve pattern recognition, and allow for a deeper understanding of the neural representations underlying consumer preferences.

Acknowledgment

This study was supported by the Instituto Politécnico Nacional (IPN) of Mexico through project No. 20250776 under the project titled *Neurodecodificación de Preferencias Alimentarias: Estrategias de Prevención Basadas en Inteligencia Artificial para la prevención de la Epidemia de Obesidad-Diabetes en México*, funded by the Secretaría de Investigación y Posgrado, Comisión de Operación y Fomento de Actividades Académicas, and by the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) of Mexico. This study is also part of the 2025 Call for InterInstitutional Collaboration Projects IPN-UAM-UAEMÉX under the project titled *Desarrollo de una Aplicación de Inteligencia Artificial para el seguimiento de contaminantes, salud, y Análisis de Factores Determinantes para el Estado de México*, by through Project No. IPCC-008-2024. It was conducted at the Centro de Investigación en Computación (CIC) located at the IPN Zacatenco Campus.

References

1. **Alianza por la Salud Alimentaria (2024).** Se podrían evitar 27,700 muertes al año por causa de enfermedades cardiovasculares si la población mexicana redujera su consumo de sodio, de acuerdo con las recomendaciones de la OMS. Consultado el 2 de abril de 2025.
2. **Avilés-López, S. I., Basurto-Pensado, M. , Palillero-Sandoval, O., Castillo-Velásquez, F. A. (2023).** Development of membranes for use as the seismic mass in a fiber optic-accelerometer. Computación y Sistemas, Vol. 27, No. 3. DOI: 10.13053/cys-27-3-4607.
3. **Barrientos-Aguilar, A., Gamboa-Cruzado, J., López de Montoya, R., Ortiz Céliz, N. M., Asto Huaman, L., Sinche Crispin, F., Ríos-Toledo, G. (2025).** Transforming freight transport with business intelligence: A case study in the peruvian logistics sector. Computación y Sistemas, Vol. 29, No. 1. DOI: 10.13053/cys-29-1-5544.

4. **Beigrezaei, S., Jambarsang, S., Khayyatzadeh, S. S., Mirzaei, M., Mehrparvar, A. H., Salehi-Abargouei, A. (2023).** A comparison of principal component analysis, partial least-squares, and reduced-rank regressions in the identification of dietary patterns associated with hypertension: YaHS-TAMYZ and Shahedieh cohort studies. Frontiers in Nutrition, Vol. 9, pp. 1076723. DOI: 10.3389/fnut.2022.1076723.

5. **Bergmann, D., Stryker, C. (2023).** ¿qué es un autocodificador? https://www.ibm.com/mx-es/think/topics/autoencoder. Consultado el 11 de agosto de 2025.

6. **Bhattacharya, J., Kanjilal, P., Nizamie, S. (2000).** Decomposition of posterior alpha rhythm. IEEE Transactions on Biomedical Engineering, Vol. 47, No. 6, pp. 738–747. DOI: 10.1109/10.844222.

7. **Born, E., Lipskaia, L., Breau, M., Houssaini, A., Beaulieu, D., Marcos, E., Pierre, R., Cruzeiro, M. D., Lefevre, M., Derumeaux, G., Bulavin, D. V., Delcroix, M., Quarck, R., Reen, V., Gil, J., Bernard, D., Flaman, J.-M., Adnot, S., Abid, S. (2023).** Eliminating senescent cells can promote pulmonary hypertension development and progression. Circulation, Vol. 147, No. 8, pp. 650–666. DOI: 10.1161/CIRCULATIONAHA.122.058794.

8. **Builtin (2025).** Step-by-step explanation of principal component analysis. https://builtin.com/data-science/step-step-explanation- principal-component-analysis.

9. **Cahyana Resky, A. A., Lapendy, J. C., Nur Risal, A. A., Surianto, D. F., Wahid, A. (2025).** Pca and t-sne implementation for knn hypertension classification visualization. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), Vol. 9, No. 1, pp. 175–184. DOI: 10.29207/resti.v9i1.6208.

10. **Ding, S., Ye, J., Hu, X., et al. (2024).** Distilling the knowledge from large-language model for health event prediction. Scientific Reports, Vol. 14, pp. 30675. DOI: 10.1038/s41598-024-75331-2.

11. **Dolu-Surabhi, S. N., Rajagopal, M., Hema Kumar, M., Srividhya, S. (2024).** Prescriptive analytics on cloud based systems using deep learning techniques. Computación y Sistemas, Vol. 28, No. 3. DOI: 10.13053/cys-28-3-5179.

12. **García Amaro, E., Cervantes Canales, J., García Lamont, F., Lara Viveros, F. M., Ruiz Castilla, J. S., Espejel Cabrera, J. (2024).** Use of computer vision techniques for recognition of diseases and pests in tomato plants. Computación y Sistemas, Vol. 28, No. 2. DOI: 10.13053/cys-28-2-3927.

13. **Gobierno de México, Secretaría de Salud (2025).** En México, 40 millones de personas presentan hipertensión arterial. Consultado el 2 de abril de 2025.

14. **Gobierno de México, Secretaría de Salud (2025).** Sector Salud emprenderá campaña para disminuir de 5 a 10% la hipertensión arterial, diabetes mellitus y obesidad. Consultado el 2 de abril de 2025.

15. **Hung, M. H., Chen, C.-H., Tseng, Y.-H., Huang, C.-C. (2024).** Abstract 4145524: Artificial intelligence-based screening for blood pressure phenotypes of white-coat and masked hypertension in outpatient settings. Circulation, Vol. 150, No. Suppl_1, pp. A4145524–A4145524. DOI: 10.1161/circ.150.suppl_1.4145524.

16. **Jimenez, C. O. S., Mesa, H. G. A., Rebolledo-Mendez, G., de Freitas, S. (2011).** Classification of cognitive states of attention and relaxation using supervised learning algorithms. 2011 IEEE International Games Innovation Conference (IGIC), pp. 31–34. DOI: 10.1109/IGIC.2011.6115125.

17. **Jr., R. D. S., Belcaro, G. V., Laurora, G., Cesarone, M. R., Sanctis, M. T. D., Incandela, L. (1998).** Ocular and orbital blood flow in patients with essential hypertension treated with trandolapril. Retina, Vol. 18, No. 6, pp. 539–545.

18. **Jusic, A., Junuzovic, I., Hujdurovic, A., Zhang, L., Vausort, M., Devaux, Y. (2023).** A

machine learning model based on micrornas for the diagnosis of essential hypertension. Non-Coding RNA, Vol. 9, No. 6, pp. 64. DOI: 10.3390/ncrna9060064.

19. **Kazemi-Díaz, A., Bote-Curiel, L., Sabater-Molina, M., Gimeno-Blanes, J. R., Sala-Pla, S., Gimeno-Blanes, F. J., Muñoz-Romero, S., Rojo-Álvarez, J. L. (2023).** An embedding approach for biomarker identification in hypertrophic cardiomyopathy. 2023 Computing in Cardiology (CinC), Vol. 50, pp. 1–4. DOI: 10.22489/CinC.2023.123.

20. **Li, K. G., Shapiai, M. I., Adam, A., Ibrahim, Z. (2016).** Feature scaling for eeg human concentration using particle swarm optimization. 2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE), pp. 1–6. DOI: 10.1109/ICITEED.2016.7863292.

21. **Liu, Y., Su, H., Li, C. (2024).** Effect of inverse solutions, connectivity measures, and node sizes on eeg source network: A simultaneous eeg study. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 32, pp. 2644–2653. DOI: 10.1109/TNSRE.2024.3430312.

22. **Liu, Z., Li, H., Li, W., Zhuang, D., Zhang, F., Ouyang, W., Wang, S., Bertolaccini, L., Alskaf, E., Pan, X. (2023).** Noncontact remote sensing of abnormal blood pressure using a deep neural network: a novel approach for hypertension screening. Quantitative Imaging in Medicine and Surgery. DOI: 10.21037/qims-23-970.

23. **López, J. F. (2024).** Covarianza: qué es, cómo se calcula y un ejemplo. https://economipedia.com/definiciones/covarianza.html. Actualizado el 22 de marzo de 2024.

24. **Mamalakis, M., Dwivedi, K., Sharkey, M., et al. (2023).** A transparent artificial intelligence framework to assess lung disease in pulmonary hypertension. Scientific Reports, Vol. 13, pp. 3812. DOI: 10.1038/s41598-023-30503-4.

25. **Nunez-Garcia, I., Lizarraga-Morales, R. A., Hernandez-Belmonte, U. H., Jimenez-Arredondo, V. H., Lopez-Alanis, A. (2022).** Classification of paintings by artistic style using color and texture features. Computación y Sistemas, Vol. 26, No. 4. DOI: 10.13053/cys-26-4-4022.

26. **Pachori, D., Gandhi, T. K. (2024).** Fbse-based approach for discriminating seizure and normal eeg signals. IEEE Sensors Letters, Vol. 8, No. 12, pp. 1–4. DOI: 10.1109/LSENS.2024.3493253.

27. **Pentikäinen, M., Simonen, P., Tuunanen, H., Leskelä, P., Harju, T., Jääskeläinen, P., Asseburg, C., Oksanen, M., Soini, E., Wennerström, C., Puhakka, A. (2024).** Treatment patterns for pulmonary arterial hypertension (pah) and chronic thromboembolic pulmonary hypertension (cteph) in finland between 2008 and 2020 - a descriptive real-world cohort study (finpah). The Journal of Heart and Lung Transplantation, Vol. 43, No. 4, pp. S407. DOI: 10.1016/j.healun.2024.02.1304.

28. **Pérez-Bueno, F., García, L., Maciá-Fernández, G., Molina, R. (2022).** Leveraging a probabilistic pca model to understand the multivariate statistical network monitoring framework for network security anomaly detection. IEEE/ACM Transactions on Networking, Vol. 30, No. 3, pp. 1217–1229. DOI: 10.1109/TNET.2021.3138536.

29. **Ríos-Toledo, G., Velázquez-Lozada, E., Posadas-Duran, J. P. F., Prado Becerra, S., Pech May, F., Monjaras Velasco, M. G. (2023).** Evaluation of feature extraction techniques in automatic authorship attribution. Computación y Sistemas, Vol. 27, No. 2. DOI: 10.13053/cys-27-2-4623.

30. **Sakuraba, S., Kobayashi, H., Sakai, S., Yokosawa, K. (2013).** Alpha-band rhythm modulation under the condition of subliminal face presentation: Meg study. 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 6909–6912. DOI: 10.1109/EMBC.2013.6611146.

31. Saurabh, S., Gupta, P. K. (2024). Detection and classification of multiple sclerosis from brain mris by using mobilenet 2d-cnn architecture. Computación y Sistemas, Vol. 28, No. 3. DOI: 10.13053/cys-28-3-4197.

32. Silva, J. A. I. R., Suarez Burgos, F. E., Wu, S.-T. (2016). Interactive visualization of the cranio-cerebral correspondences for 10/20, 10/10 and 10/5 systems. 2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), pp. 424–431. DOI: 10.1109/SIBGRAPI.2016.065.

33. Singh, N., Eickhoff, C., Garcia-Agundez, A., et al. (2023). Transcriptional profiles of pulmonary artery endothelial cells in pulmonary hypertension. Scientific Reports, Vol. 13, pp. 22534. DOI: 10.1038/s41598-023-48077-6.

34. t-SNE, **.** t-SNE (T-distributed Stochastic Neighbor Embedding). https://interactivechaos.com/es/manual/tutorial-de-machine-learning/t-sne, year = 2025,.

35. UMAP (2025). UMAP: Uniform manifold approximation and projection for dimension reduction. https://umap-learn.readthedocs.io/en/latest/.

36. UNAM Global (2025). Dieta saludable, clave para prevenir la hipertensión. Consultado el 2 de abril de 2025.

37. UNAM Global (2025). Hipertensión afecta a 30% de mexicanos y eleva costos en salud. Consultado el 2 de abril de 2025.

38. Wen, Z., Jiang, L., Yu, F., Xu, X., Chen, M., Xue, J., Zhu, P., Ying, Z., Li, Z., Chen, T. (2023). Single-cell landscape reveals nampt mediated macrophage polarization that regulate smooth muscle cell phenotypic switch in pulmonary arterial hypertension. DOI: https://doi.org/10.1101/2023.07.04.547668.

39. Wu, L., Zhang, W., Li, S., Li, Y., Yuan, Y., Huang, L., Cao, T., Fan, L., Chen, J., Wang, J., Liu, T., Wang, J. (2023). Transcranial alternating current stimulation improves memory function in alzheimer's mice by ameliorating abnormal gamma oscillation. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 31, pp. 2060–2068. DOI: 10.1109/TNSRE.2023.3265378.

40. Yadav, K., Tanwar, N., Hasija, Y. (2024). Exploring the role of gut microbiota in hypertension: Insights from machine learning and explainable ai. 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1–5. DOI: 10.1109/ICCCNT61001.2024.10725123.

41. Yamamoto, S., Oiwa, K., Nanai, Y., Nagumo, K., Nozawa, A. (2024). Detección de hipertensión en imágenes faciales de bandas visibles e infrarrojas cercanas mediante codificación dispersa. IEEJ Transactions on Electrical and Electronic Engineering, Vol. 144, No. 7, pp. 672–678. DOI: 10.1541/ieejeiss.144.672.

42. Ye, H., Goerttler, S., He, F. (2024). Eeggmacn: Interpretable eeg graph mutual attention convolutional network. 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1–4. DOI: 10.1109/EMBC53108.2024.10782694.

43. Zaldivar-Herrera, J. E., Sánchez-Fernández, L. P., Rodríguez-Méndez, L. M., Zagaceta-Álvarez, M. T. (2024). A fuzzy inference model for evaluating data transfer in lte mobile networks via crowdsourced data. Computación y Sistemas, Vol. 28, No. 3. DOI: 10.13053/cys-28-3-4826.

44. Zaldívar, H. M., Posadas, C., Manero, L., Torres, G. S. (1986). Evaluation of sympathetic tone in essential hypertension by the clonidine suppression test. Archivos del Instituto de Cardiología de México, Vol. 56, No. 3, pp. 211–217.

45. Zhao, J., Guo, W., Wang, J., Wang, T. (2023). Exploring the association of dietary patterns with the risk of hypertension using principal balances analysis and principal component analysis. Public Health Nutrition,

Vol. 26, No. 1, pp. 160–170. DOI: 10.1017/S136898002200091X.

46. Zhong, C., Si, Y., Yang, H., Zhou, C., Chen, Y., Wang, C., Liu, Y., Chen, C., Shi, H., Lai, X., Tang, H. (2023). Identification of monocyte-associated pathways participated in the pathogenesis of pulmonary arterial hypertension based on omics-data. Pulmonary Circulation. DOI: 10.1002/pul2.12319.

Article received on 12/08/2025; accepted on 12/09/2025.
**Corresponding author is Jesús Jaime Moreno Escobar.*