UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Revista Científica

http://revistas.udistrital.edu.co/ojs/index.php/revcie/index

ARTÍCULO DE REVISIÓN

Analysis of ML Models for the Recognition of 3D Objects Based on Biometric Eye-Tracking Data in the Context of UX and User-Centered Design

Análisis de modelos de ML para el reconocimiento de objetos 3D con base en datos biométricos de *eye-tracking* en el contexto de la UX y el diseño centrado en el usuario

Análise de modelos de ML para o reconhecimento de objetos 3D com base em dados biométricos de *eye-tracking* no contexto da UX e do *design* centrado no usuário

Angel Eduardo Villegas Ortíz[1]
Francisco Javier Rodríguez López[2]
Eduardo Emmanuel Rodríguez López[3]

Recibido: 17 de diciembre de 2025 **Aceptado:** 30 de enero de 2026

Para citar este artículo: Villegas, A. E., Rodríguez, F. J. y Rodríguez, E. E. (2026). Analysis of ML Models for the Recognition of 3D Objects Based on Biometric Eye-Tracking Data in the Context of UX and User-Centered Design. *Revista Científica*, *53*(1), 1-18. https://doi.org/10.14483/23448350.24601

Abstract

Recent advances in the analysis of machine learning (ML) models for accurate prediction have led to new approaches for classifying and clustering variables related to user experience (UX). In this study, however, the focus is not only on modeling variables, but also on capturing the dynamics of visual attention in relation to object properties. We evaluate models such as KNN, SVM, and K-means, which emerge as the most relevant for classifying visual attention patterns. When combined with the principles of user-centered design, this approach enables the construction of user profiles based on their interaction capabilities with the interface. Furthermore, it examines how 3D object properties—such as shape, color, shading, direction of motion, and motion acceleration—influence the information conveyed to users, thus shaping their UX directly. This method aims to enhance UX design and enable more precise representation of target users. In 11 exploration tests, the three ML models analyzed obtained an accuracy of 44, 72, and 65%.

Keywords: UCD, eye-tracking, user profile design, machine learning, 3D objects

1. Universidad Autónoma de Aguascalientes, Aguascalientes, Ags., México. Contacto, angel.villegas@edu.uaa.mx
2. Universidad Autónoma de Aguascalientes, Aguascalientes, Ags., México. Contacto, fjalvar@correo.uaa.mx
3. Universidad de Monterrey, San Pedro Garza García, Nuevo León, México. Contacto, eduardoemmanuel.rodriguez@udem.edu

Resumen

Los avances recientes en el análisis de modelos de aprendizaje automático (ML) para la predicción precisa han dado lugar a nuevos enfoques para clasificar y agrupar variables relacionadas con la experiencia de usuario (UX). En este estudio, sin embargo, el enfoque no se limita a modelar variables, sino también a capturar la dinámica de la atención visual en relación con las propiedades de los objetos. Se evalúan modelos como KNN, SVM y K-means, que se perfilan como los más relevantes para la clasificación de patrones de atención visual. Al combinarse con los principios del diseño centrado en el usuario, este enfoque permite la construcción de perfiles de usuario basados en sus capacidades de interacción con la interfaz. Además, se examina cómo las propiedades de los objetos 3D—como la forma, el color, el sombreado, la dirección del movimiento y la aceleración del movimiento—influyen en la información transmitida a los usuarios, configurando así su UX de manera directa. Este método tiene como objetivo mejorar el diseño de la UX y permitir una representación más precisa de los usuarios objetivo. En 11 pruebas exploratorias, los tres modelos de ML analizados obtuvieron precisiones del 44, 72 y 65 %.
Palabras clave: UCD, seguimiento ocular, diseño de perfiles de usuario, aprendizaje automático, objetos 3D

Resumo

Avanços recentes na análise de modelos de aprendizado de máquina (ML) para previsão precisa têm levado a novas abordagens para classificar e agrupar variáveis relacionadas à experiência do usuário (UX). Neste estudo, no entanto, o foco não se limita à modelagem de variáveis, mas também à captura da dinâmica da atenção visual em relação às propriedades dos objetos. Avaliam-se modelos como KNN, SVM e K-means, que se destacam como os mais relevantes para a classificação de padrões de atenção visual. Quando combinada com os princípios do design centrado no usuário, essa abordagem permite a construção de perfis de usuário com base em suas capacidades de interação com a interface. Além disso, examina-se como as propriedades dos objetos 3D—como forma, cor, sombreamento, direção do movimento e aceleração do movimento—influenciam a informação transmitida aos usuários, moldando diretamente sua UX. Esse método tem como objetivo aprimorar o design de UX e permitir uma representação mais precisa dos usuários-alvo. Em 11 testes exploratórios, os três modelos de ML analisados obtiveram acurácias de 44, 72 e 65%.
Palavras-chaves: UCD, rastreamento ocular, design de perfis de usuário, aprendizado de máquina, objetos 3D

INTRODUCTIon

As user experiences become increasingly complex to design, implement, and evaluate, paradigms such as user-centered design (UCD) advocate for crafting experiences that prioritize user satisfaction by creating representative profiles of potential software users. This paper explores how the integration of biometric data—specifically obtained through eye-tracking—opens new avenues for developing such user profiles.

This article begins by presenting a rationale for the proposed approach, followed by a section outlining key concepts that support and contextualize the initial methodological proposals. Finally, it examines a case study and its preliminary results, concluding with a reflection on future work and a discussion of the proposed model's potential and validity.

JUSTIFICATION

The current landscape of human-computer interaction (HCI) highlights a growing demand for integrating emerging technologies to enhance and refine its implementations (Gorichanaz, 2024; Görücü *et al.*, 2025; Uusitalo *et al.*, 2024). Tools and methodologies that enable the acquisition of biometric data from users are gaining relevance, as such data can significantly enrich the analysis of user experience (UX).

Among these, eye-tracking stands out as a non-invasive technique capable of capturing meaningful user responses to stimuli without disrupting UX interaction flows (del Punta *et al.*, 2024).

Despite its potential, the study of three-dimensional experiences remains relatively unexplored (Bruce & Hawes, 2015; Romanowski, 2019; Téllez Rojas *et al.*, 2021). This gap presents a valuable opportunity to investigate how human visual attention behaves when interacting with 3D objects. The existing literature reveals substantial differences between the observation of two-dimensional (2D) and three-dimensional (3D) stimuli. As a result, many established UX conventions, paradigms, metrics, and variables developed for 2D contexts cannot be directly applied to 3D environments due to the pronounced divergence in visual attention (VA) mechanisms (Bruce & Hawes, 2015; Kumar *et al.*, 2024; Sun *et al.*, 2023).

OBJECTIVES

With the purpose of offering valid mechanisms for UX analysis via machine learning (ML) models, the following objectives were set as guidelines for representing results and monitoring methodological proposals.

General objective

– To objectively identify users who exhibit a recognizable VA pattern across different algorithms, throughout the various stages and stimuli presented during a test, using strict biometric eye-tracking (ET) data.

Specific objectives

– To propose a set of stimuli that elicit VA responses measurable through ET technologies and analyzable using ML techniques.

- To conduct valid experimental executions and process ET trials in order to obtain a dataset that is suitable for experimentation.
- To apply ML techniques to the data obtained via ET to derive convergence metrics of VA patterns and provide interpretations of the results.

KEY CONCEPTS

In order to develop valid proposals with broad applicability across a wide range of projects, the following set of concepts is presented. These serve as the foundation for the various methodological approaches discussed in this research work.

User-centered design

UCD is a paradigm within software development processes that actively involve potential users in shaping the product. Among the advantages of this paradigm is the ability to produce software with strong performance in terms of utility and usability for a well-defined group of users (Nagro & Aldekhail, 2021).

The method employed to identify different types of users and the interaction properties that the product must possess to be truly useful and usable involves abstracting potential users into profiles. This profiling process encapsulates demographic data, prior knowledge in specific domains, computing skill levels, product preferences, or even ideological tendencies, among other characteristics. These elements are used to establish challenges, strengths, desirable features, and other aspects relevant to UX design (Calderón Ynoñan *et al.*, 2022; Chochoiek, 2017).

Conventionally, however, a profile designed within UCD may not adequately reflect real-world users. The reason is straightforward: as profile designers, it is easy to generate inaccurate assumptions about user behavior, and, in many cases, it is impossible to fully determine the effects of design properties on users (Eke *et al.*, 2019). This uncertainty in profiling suggests that, while the UCD model yields positive outcomes, it still has considerable room for improvement.

Eye-tracking

ET techniques enable the collection of biometric data related to a person's VA when observing a stimulus. Although each individual may exhibit unique behaviors based on personal preferences or interests, cognitive and VA processes govern the scanning patterns of a stimulus, shaped by prior experience and the basic properties of the stimulus (Bednarik & Tukiainen, 2006; Duchowski, 2017). These patterns can be analyzed through ET, providing a valuable, precise, and—most importantly—objective source of information about a test subject.

A VA process encompasses several specific ocular behaviors. These include, among others, object tracking, reading, silhouette and face detection, and the estimation of size and distance. Each of these generates a specific set of ocular reactions that an eye-tracker can easily parameterize (Madi, 2025).

A central component of VA analysis involves fixations and saccadic movements, which can be further extended into more precise metrics. A fixation is defined as the retention of a person's optical axis on the same point for approximately 250 ms (according to the metric provided by Tobii Technologies). In most cases, a single fixation alone does not provide extensive information about human vision. Therefore, the analysis also considers saccadic movements, which are the transitions between one fixation and another (Albiz *et al.*, 2023).

When used together throughout a test, these two variables can adequately parameterize the VA processes performed by the subject during the experiment. Thanks to modern ET technologies, it is possible to obtain precise and filtered information in the form of metrics derived from experimental executions.

Visual attention. VA is understood as the process of obtaining information from a stimulus, interpreting it, and subsequently reacting to it through various subprocesses such as focusing, ET, distance estimation, and the perception of properties like color and shape (del Punta *et al.*, 2024; Hoffman & Subramaniam, 1995; Tarazona Evangelista *et al.*, 2018). To this effect, VA draws on a diversity of cognitive processes that allow utilizing an individual's prior experience to understand and interpret a stimulus as a function of the object's basic properties while relying on the context and the information available at the time (Bisley, 2011; Ferreres, 2023; Rueda Cuerva, 2021).

As previously mentioned, the process of analyzing VA differs between 2D and 3D elements, as each involves distinct patterns and considerations. The most fundamental characteristic in interpreting a 3D object—beyond its shape, texture, size, and distance—is the shadow it projects. Shadows allow the observer to distinguish features such as movement, acceleration, or direction of displacement within a defined environment (Besl & Jain, 1985; Chelloug *et al.*, 2022; Liu *et al.*, 2022).

Convergent and divergent patterns. To delve into the theory of convergent and divergent patterns, it is necessary to reference prior work that relates these concepts, derived from ET to UX. In Villegas Ortíz *et al.* (2024), a comprehensive framework is described for analyzing both visual and numerical results, enabling the interpretation of user behavior in 2D experiences.

In this work, a *convergent usage pattern* is defined as any set of trials that share similarities among different users. A convergent pattern follows the principle that "any recurrence of VA patterns on the same interface indicates the natural flow of user interaction during UX" (Villegas Ortíz *et al.*, 2024. p. 101). In other words, the greater the number of users converging in their patterns, the higher the likelihood that the interface will demonstrate strong effectiveness, clarity, efficiency, and user satisfaction.

Conversely, a *divergent pattern* refers to any behavior that does not align with most of the population. Far from being disregarded, divergent patterns are proposed as the most relevant profiles to consider, since incorporating a greater variety of divergent usage patterns into UX design results in more robust and usable interfaces for a broader range of users.

Furthermore, in Villegas Ortíz *et al.* (2024), visual results largely determined the capacity to analyze convergent *vs.* divergent patterns. However, this study aims to construct mechanisms for evaluating these concepts in a numerical and objective manner through the implementation of ML models.

Machine learning. According to Leivi (2019), ML can be defined as "a type of AI that enables machines to learn directly from examples and experience, which is acquired from a dataset through training" (p. 26). This refers to the processes required to provide an input dataset and obtain outputs based on the patterns identified within the data.

In relation to the above, it is assumed that VA patterns are sufficiently distinguishable from one another to allow for predictions from ET trials. To this effect, two branches of artificial intelligence (AI) can be employed to model the problem and conduct an initial exploration of the capacity of ET data to recognize different users through simple stimuli.

Supervised learning by classification. This technique enables the classification of data, either in binary or multiclass form. The structure shared by supervised learning algorithms divides the dataset into two subsets: training variables and testing variables (train-test split). Each subset comprises two components, denoted as X and y, where X represents the data that make up the dataset, while y—referred to as the *predictor variable* or *target*—represents the original class and the predictions assigned by the ML model (Ali *et al.*, 2020; Thakur, 2020).

Unsupervised learning. The main characteristic of unsupervised learning models is the absence of a predictor variable or target. In other words, unsupervised learning models do not classify data according to predefined classes. Instead, similarities among data patterns allow grouping the results (commonly referred as *clustering* or *clusters*) (Kashtalian & Sochor, 2021).

Through these two categories of ML, the problem can be approached from two perspectives. Unsupervised learning offers categorization without relying on a predefined target variable, allowing for groupings that are solely based on the data. From a more conservative UCD perspective, one could predetermine the number of groups to align with the model's results. For instance, if only four profiles are expected, this parameter could be set as the value of K in clustering. However, in the works discussed below, the number and types of profiles are not defined by the designer's criteria; the unsupervised model's results are expected to define the possible profiles based on their objective usage

Profiling through ET biometric data

During the 19th century, many sciences and fields of study began to gain popularity and develop models supported by the scientific method and the concept of theoretical validity. One of the disciplines that flourished during this period was psychology, which broke away from many established conventions to propose new methods for the objective study of the human psyche (Burgos, 2017). As a method, profiling emerged in the early 20th century as a tool that standardized psychological testing across patients, leading to significant advances in the diagnosis, classification, and study of mental illnesses, as well as in understanding the relationship between behavioral patterns and the psyche, among other contributions (Burgos, 2017; Hernández Salazar, 1993).

In the second half of the 20th century, a new type of profiling analysis arose in response to needs within the field of criminology. Today, criminal profiling remains one of the most established and successful methodological applications in psychology (Jiménez *et al.*, 2012). This work draws on some of the fundamental ideas from that field, although its purpose is very different from the original concept.

Building on the idea of generating a process that allows categorizing individuals according to their 'behavior', the concept of *profiling* proposed in this study refers to the overall influence of VA patterns in the identification of convergent patterns (Wu *et al.*, 2024).

Earlier, the design of profiles within UCD was discussed as a cornerstone of the abstraction process for user characteristics in the design of potential user representations. In recent years, this concept has enabled the development of more inclusive and equitable interfaces and experiences (Görücü *et al.*, 2025).

With this approach, beyond proposing a particular analysis of ET data when interacting with a specific set of stimuli, the aim is to establish a precedent for implementing these technologies in the design of profiles based on biometric data.

METHODOLOGICAL PROPOSALS

User profiling through supervised and unsupervised learning models represents an extensive process that requires the integration of multiple. This section provides a detailed description of the steps established for conducting analysis of the convergent patterns observed across a group of test subjects.

Stimuli selection. The selection of stimuli is crucial for eliciting differentiable VA patterns that allow classifying each individual appropriately. While the concepts of *convergent* and *divergent patterns* have already been introduced, the following question remains: which properties of the stimuli are responsible for producing more specific patterns?

In the case of 3D objects, this work proposes two sets of characteristics that maximize perceptual differences in human vision when exposed to 3D stimuli. The first group refers to the intrinsic properties of an object which allow the user to identify it at a given instant t throughout experience. The second group refers to the displacement properties of the object within an environment and their relationship with cognitive-visual processes that enable the prediction of trajectories or the continuous localization of the object on a screen (Tables 1 and 2).

Table 1. *Intrinsic properties of a 3D object*

Property	Description	Reference
Distribution (location)	The user observes and compares the environment to infer the approximate position and distance of the object.	(Liu *et al.*, 2022)
Shape and color	Initial guesses made by an observer after spatially locating the object.	(Liu *et al.*, 2022; Ustárroz & Grandi, 2016)
Instantaneous size relative to the observer	Alongside the identification of the object's silhouette, the observer estimates its size based on the type of object and additional cues, such as the shadow it projects.	(Chelloug *et al.*, 2022)
Fixed shadow	In many cases, to estimate the size of an object, the brain relies on the shadow it casts. This is particularly relevant in digital 3D environments, where shadows allow the observer to spatially locate the object and maximize the use of the available information.	(Correa Torres, 2021)

Table 2. *Displacement properties of a 3D object*

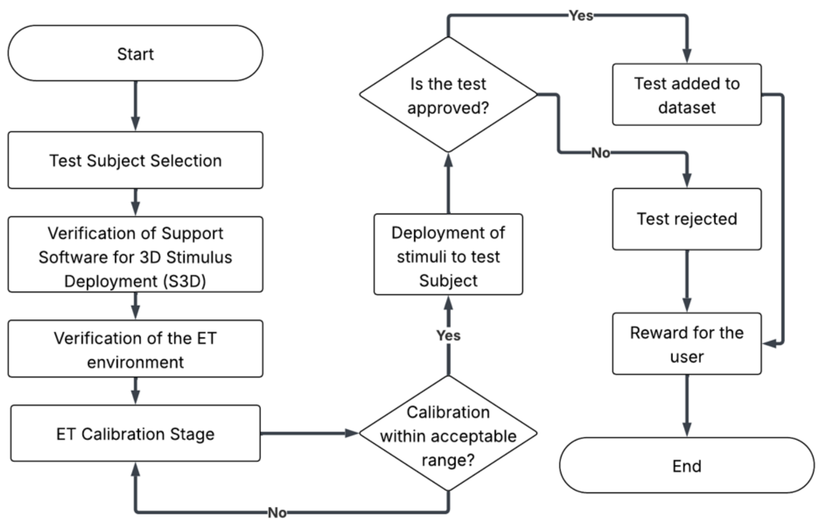
Property	Description	Reference
Reading pattern in motion	One of the most frequently analyzed patterns through ET is reading. However, reading on a static screen is entirely different from reading while the text is in motion.	(Madi, 2025; Mondal *et al.*, 2022)
Speed or acceleration relative to the observer	Tracking is one of the most complex VA processes. Cognitive-visual mechanisms allow for precise estimations of acceleration and speed. If the object moves within an appropriate optical range, a person can follow objects at high acceleration and even perceive details such as silhouette, color, and shape with moderate success.	(Rueda Cuerva, 2021)
Rotation	A less intuitive type of movement that rarely occurs in the real world. This behavior triggers specific VA reactions.	(Sun *et al.*, 2023)

Experimental design. To obtain valid and highly precise biometric data, it is necessary to establish an experimental protocol that allows the team to generate robust trials while minimizing the need to discard them. Based on extensive experience with experimental ET executions, the importance of proper calibration is emphasized. As shown in Figure 1, the only iterative cycle in the process occurs during the calibration stage.

Due to the execution conditions of the Unity graphics engine, a decision was made to use the Build tool, which generates a package containing the project executable, thereby reducing computational resource consumption. In this case, the final version of the build was named *Support Software for 3D Stimulus Deployment* (S3D).

For discarding trials, calibration results were primarily considered. The threshold for a valid trial was determined by the distance between a calibration point on the screen and the user's fixation, with a maximum deviation of 35 pixels.

Figure 1. *Experimental protocol diagram*



Processing of ET trials and dataset generation. An experimental ET deployment provides several elements: calibration data, a test and user logbook, a recording of the interface or stimuli analyzed, and some automatic and control metrics, among others.

To produce a valid trial, two main processes are carried out: (i) locating the stimuli in time and (ii) locating the stimuli spatially on the screen. In the first case, markers are placed along the recording to determine the exact moment when a stimulus is displayed or when a particular action occurs for analysis.

These events, called times of interest (TOI), may correspond to specific actions within UX. For example, when analyzing an e-commerce store, an event could be adding a product to the cart or completing the online purchase process. Properly defining these events is essential for accurate ET analysis.

Similarly, areas of interest (AOI) establish the boundaries of the regions where gaze fixations are expected to cluster on the screen during the trial. These AOI may include buttons, images, text blocks, standalone elements, or even advertising banners. In essence, within UX analysis, an AOI is associated with the specific elements of the interface under study.

In addition to these variables, the interval is understood as the identifier of distinct repeated sub-experiences. When TOI or AOI are repeated across records, the interval is used to separate each of these groups. It should be noted that all configurations of these parameters are determined at the discretion of the designer, since the tools used in this work do not allow for automatic trial processing.

Once TOI and AOI had been assigned to each trial, and intervals and other control variables such as calibration results had been reviewed, a dataset was generated using Tobii Pro Lab's Metrics Export tool. This tool allows filtering ET variables and metrics for valid trials.

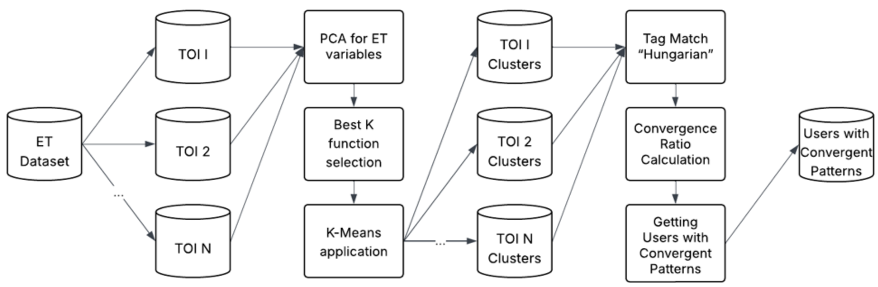
Profile processing through ML. The profiling analysis proposed in this work employs two processes: supervised and unsupervised learning. Although there are some differences in the grouping mechanisms, both processes consume the same dataset and output the set of users with convergent VA patterns.

Unsupervised learning model. In unsupervised learning, the grouping variable is defined for each TOI. Once the original dataset is separated by TOI, principal components analysis (PCA) is applied to reduce the number of ET variables, resulting in a more compact dataset.

The BestK function performs an elbow analysis to obtain a value of *K* for each dataframe derived from the TOI. It then averages the best individual *K* values and selects a global *K*, which is applied uniformly to each dataframe. Finally, clustering is performed using K-means, and clusters are obtained for each user.

The tag match process is used to relate the clusters of each individual TOI to one another, since they will be compared in the next step, called convergence ratio calculation. This step estimates the number of recurring clusters across the TOI for each person. Figure 2 illustrates this process.

Figure 2. *Unsupervised learning model diagram*



Supervised learning model. In this research, a different grouping variable was selected for comparison with the unsupervised learning model. The interval was used as the variable that enables the prediction of participant labels (Figure 3). The application of PCA to reduce ET variables was maintained for the interval dataframes. However, to explore the properties of this type of ML, two alternatives were considered: the compact and the detailed dataframe.

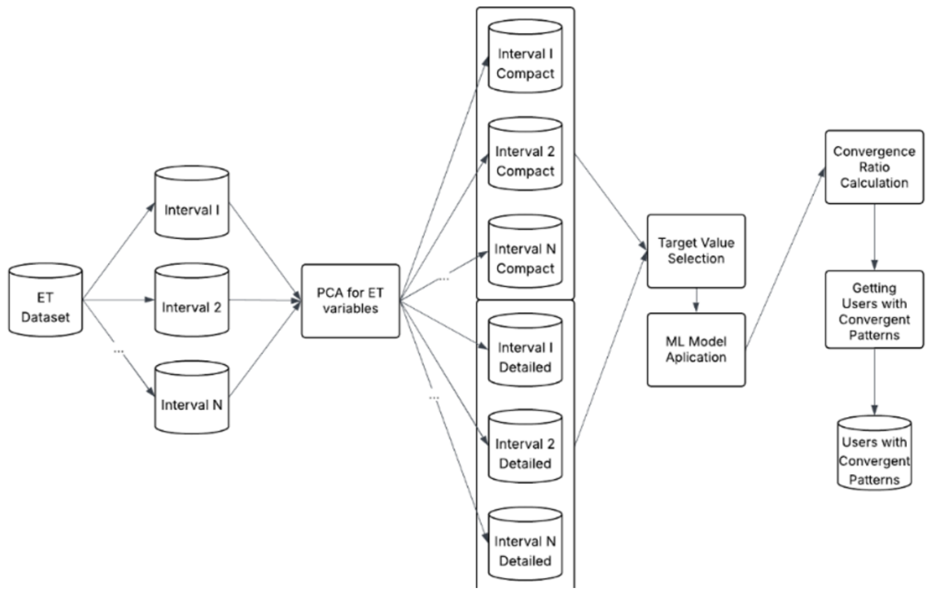
The compact dataframe aggregates, through an additional PCA process, the recurring events within the intervals. Each interval refers to a set of TOI that are repeated sequentially throughout a trial. Therefore, there will be n records per person for each TOI established in each interval dataframe. The compact dataframe consolidates these multiple records of a person into a single entry while preserving the previously obtained PCA columns of ET.

In contrast, the detailed dataframe retains each of the multiplied records for every person and every TOI generated. In both cases, a target label is assigned to the participant, and, for this model, the calculation of convergence ratios is performed using Equation (1).

$$\begin{gathered}\textit{Convergence Ratio for Interval 1} = \\ \textit{train\_X=Interval 1, test\_X = Universe\_Intervals} - \{\textit{Interval1}\}\end{gathered} \tag{1}$$

Applying the concepts of supervised learning, this Equation uses a pivot interval dataframe as *X_True*, to which the model is applied. Once the model is trained with the pivot interval, labels are generated for the remaining intervals of the universe. The purpose of this analysis is to identify intervals that behave atypically when compared to the others. The higher the convergence ratio, the stronger the relationship across the entire UX, resulting in a convergent pattern. Conversely, a low value indicates that the interval exhibits a different behavior due to the specific properties of the TOI associated with it across different users.

Figure 3. Supervised learning model diagram.



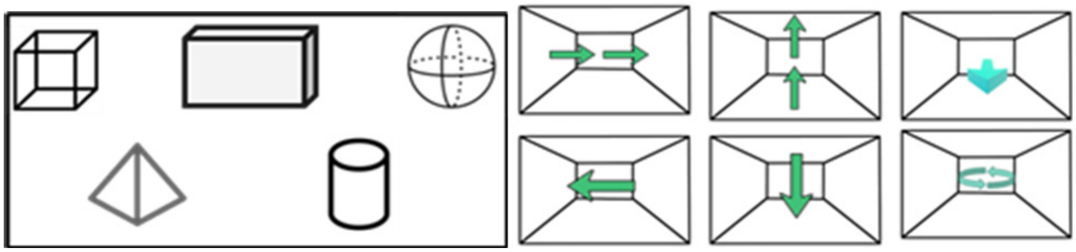
CASE STUDY

As mentioned in previous sections, to adequately analyze VA patterns, it is necessary to elicit very specific visual reactions from users in order to achieve more accurate clustering/classification. The following subsections detail the characteristics of the case study.

Experimental design. In this case, the experimental design sought to identify the threshold of VA elicited by a set of elementary 3D objects with a particular feature. In addition to moving along different trajectories (which already generates interesting VA patterns for analysis), these objects also contain text within them. Users are not informed about the directions of movement or the behavior of the objects. Instead, they are instructed to read the text contained in each element.

This reading-in-motion pattern, applied to different 3D objects moving in various directions, is expected to produce highly distinctive VA patterns. Figure 4 shows the set of 3D objects (five) and their displacement trajectories (four).

Figure 4. *Set of 3D objects and their displacement trajectories*



Processing of ET trials and dataset generation. For this case study, 13 experimental deployments were generated, out of which two were discarded due to calibration issues. In the end, ten users (grouped in the *Recording* column) were included in the final dataset. Each trial lasted no longer than three minutes, which made the overall production process particularly time-consuming.

As mentioned in the methodology section, it is necessary to locate the stimuli of interest both temporally and spatially within the trial. Figure 5 shows the process of separating each TOI. In this case, the TOI categories are related to the displacements of the object, while each interval corresponds to one of the five objects displayed during the trial. The method consists of reviewing the recording and placing markers at the time points where each TOI begins and ends.

Finally, as shown in Figure 6, AOI are delimited in the regions where visualizations and VA processes are expected to occur. By extension, the interval column refers to each 3D object.

Figure 5. *Processing of TOI for a single user trial*

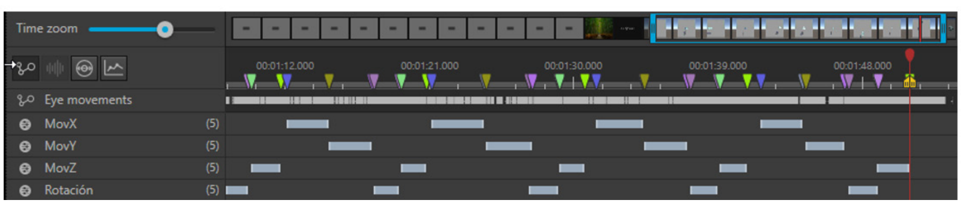
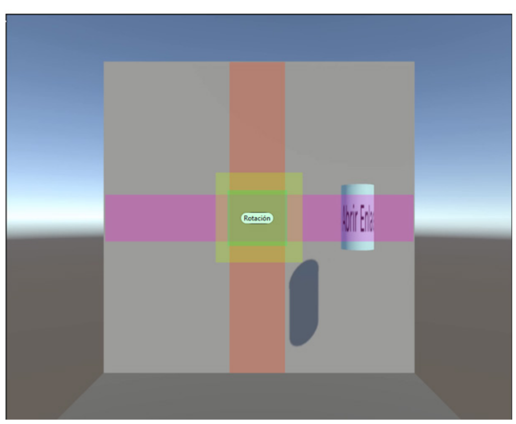


Figure 6. *Processing of AOI for a single user*



Data processing via ML. For both supervised and unsupervised learning, processing shared the same input dataset and generated results based on the similarity of the clusters/classifications obtained.

For unsupervised learning, the grouping variable was movement trajectory, indicating whether the reading-in-motion patterns were sufficiently differentiable to produce clusters of users throughout the experience.

In contrast, the supervised learning model used the interval as the grouping variable. In this case, the objective was to estimate which 3D object appeared to yield weaker relationships when compared to the others. Considering that the goal was to obtain a set of convergent users, the mechanism used to represent convergence was the recurrence of the cluster/classification throughout the experience. These results are summarized in the Figures presented in the next section.

RESULTS

This section describes the results for both procedures of VA pattern analysis applied to the designed case study. It should be noted that the final outcome was the identification of a set of convergent users, uniquely represented in the *Recording* column.

K-means model. The K-means model generates a cluster for each TOI (since TOI is the grouping variable in this process) throughout the experience. In this case, all objects marked with interval 1 are grouped together into a single record for the user. Once the TOI dataframes have been clustered, the Hungarian method is applied to maintain a consistent cluster structure across different dataframes. These dataframes, with their generated clusters, are then compared in the convergence analysis table.

The result, expressed as the convergence index ratio, refers to the algorithm's ability to recognize the same subject during the trial (Figure 7).

Figure 7. *Results for the ten experimental executions and their convergence index ratio for each displacement trajectory using the K-means model*

Índice	Recording	Cluster_X	Cluster_aligned_Y	Cluster_aligned_Z	Cluster_aligned_R	ncordant	convergence_ratio
0	Recording3	6	6	6	6	True	1
1	Recording8	1	0	1	7	False	0.5
2	Recording10	7	7	7	6	False	0.75
3	Recording1	3	2	5	3	False	0.5
4	Recording2	4	4	4	4	True	1
5	Recording9	3	3	3	1	False	0.75
6	Recording5	5	5	1	5	False	0.75
7	Recording12	2	1	2	2	False	0.75
8	Recording7	8	8	8	8	True	1
9	Recording11	1	1	4	1	False	0.75
10	Recording13	0	0	0	0	True	1

In this case, the visual attention patterns were easily recognizable for the algorithm, since four users achieved a perfect convergence ratio (CR), five obtained 75% accuracy in the clusters, and only two users showed moderate CR values (50%).

KNN model. For the supervised learning models, two experimental dataframes were considered: compact and detailed. The compact dataframe aggregated all TOI into a single record while preserving the PCA columns, whereas the detailed one retained multiple records for each participant.

In both cases, Equation (1) was applied to obtain the predictions for each interval. Figure 8 shows the results for the compact dataframe. Since there is only one record per participant for each interval, if the individual classification of the label does not match, the CR value is reduced.

Figure 8. *Results of the compact dataframe for the KNN model*

Recording	I1	I2	I3	I4	I5	%
Recording1	F	F	F	F	F	0
Recording10	F	F	F	T	T	40
Recording11	F	F	F	F	T	20
Recording12	F	F	F	F	F	0
Recording13	T	F	F	F	F	20
Recording2	F	F	F	F	F	0
Recording3	F	F	T	F	F	20
Recording5	F	F	F	F	F	0
Recording7	F	F	T	F	F	20
Recording8	F	F	F	F	F	0
Recording9	F	F	F	F	T	20

The case of the detailed dataframe allows for a more flexible analysis since, with four records per user, if at least one of them was correctly labeled, it was considered a valid convergence (Figure 9).

In this case, for the compact dataframe, only one user achieved two matches (the maximum). On the other hand, in the results for the detailed dataframe, a single record reached a perfect CR, and three users exceeded a CR of 50%.

Figure 9. *Results of the detailed dataframe for the KNN model*

Recording	I1	I2	I3	I4	I5	%
Recording1	T	T	F	F	T	60
Recording10	T	T	T	T	T	100
Recording11	T	T	F	F	T	60
Recording12	F	F	T	F	T	40
Recording13	F	F	F	F	T	20
Recording2	F	T	T	T	F	60
Recording3	F	F	F	T	F	20
Recording5	F	F	T	F	F	20
Recording7	T	T	F	F	F	40
Recording8	T	F	F	F	T	40
Recording9	F	F	F	F	F	0

SVM model. The SVM model shares the same functioning as KNN, so the classification differences lie exclusively in the properties of the model. As shown in Figure 10, the number of users exceeding the 50%

CR threshold is the same. However, on average, the predictions achieved at least one correct match in most of the records for the compact dataframe.

In contrast, the results of the detailed dataframe in Figure 11 show a substantial improvement, as it labels three users with a perfect CR and another three surpass the 50% threshold.

Figure 10. *Results of the compact dataframe for the SVM model*

Recording	I1	I2	I3	I4	I5	%
Recording1	F	F	F	T	F	20
Recording10	F	F	F	T	T	40
Recording11	F	F	F	F	F	0
Recording12	F	F	F	T	F	20
Recording13	F	F	F	T	F	20
Recording2	F	F	T	F	F	20
Recording3	F	F	F	F	F	0
Recording5	F	F	F	F	T	20
Recording7	T	T	F	F	F	40
Recording8	F	T	T	F	T	60
Recording9	F	F	F	T	F	20

Figure 11. *Results of the detailed dataframe for the SVM model*

Recording	I1	I2	I3	I4	I5	%
Recording1	T	T	T	T	T	100
Recording10	T	T	F	F	F	40
Recording11	F	T	F	T	F	40
Recording12	F	F	T	T	F	40
Recording13	F	T	F	F	T	40
Recording2	T	F	F	T	F	40
Recording3	T	F	T	T	F	60
Recording5	T	T	F	F	T	60
Recording7	T	T	T	T	T	100
Recording8	T	T	T	T	T	100
Recording9	T	T	F	F	T	60

CONCLUSIONS AND FUTURE WORK

As can be observed in the results section, the K-Means model exhibits a better fit with the proposed profiling concept. Although the models trained with supervised learning approaches show poor performance when each participant is grouped into a single record, the analysis of the detailed dataframe is more flexible than that of its compact counterpart. Thus, while the results are valid for estimating the convergence of patterns, the margin of error increases considerably.

The analysis of ML models applied to biometric data poses several challenges, particularly regarding the structure required for the initial dataset and in the preprocessing steps needed to filter it until a data structure is achieved which truly reflects the characteristics of the studied UX. In this case, trajectories and 3D objects were analyzed in detail. These properties of displacement, shape, and dimension of the figures, while providing relevant information about UX, remain very elementary. In other words, this first case study paves the way for experimentation with much more robust systems that are not limited to fixed trajectories or embedded text within objects to force user VA.

If this analysis is extended, it will soon be possible to apply ML models to data obtained from user testing with biometric inputs, in order to assess the capacity of an interface—or a set of interfaces—to be used effectively.

Building on the above, the relevance of pattern convergence lies in the fact that every designer seeks to ensure that their interfaces achieve properties allowing for the greatest number of users to follow a defined interaction flow. Once this behavior is reflected in the majority of users, the designer can confidently affirm that the design works for the target audience or a significant portion of it. The analysis of CR provides an objective and quantitative measure of the recurrence of user behaviors when interacting with a product.

ACKNOWLEDGMENTS

We would like to express our gratitude to the Autonomous University of Aguascalientes for the tools, infrastructure, and human and academic resources that enable the production of work of high scientific relevance.

As the lead author, I would like to acknowledge SECIHTI, as its “National Graduate Studies Support Grant” program made this research possible.

I would also like to congratulate María Fernanda Barrón Álvarez, a member of the research team, and all the study participants.

AUTHORS' CONTRIBUTIONS

Author	**Contribution**
Angel Eduardo Villegas Ortíz	Methodological Approach. Experimental Design. Experimental implementation and generation of Eye-tracking Data. Application of Machine Learning Models. Content Writing and Editing.
Francisco Javier Álvarez Rodríguez	Validation and refinement of methodological and experimental models. Content review and editorial revisions. Validation of results and verification of the scientific relevance of the final product.
Eduardo Emmanuel Rodríguez López	Proposal for the application of machine learning models. Validation of the application of machine learning models. Review of machine learning results. Editing and style corrections.

DECLARATION OF THE USE OF ARTIFICIAL INTELLIGENCE

The authors affirm that:

None of the methodological or experimental processes were designed or validated using AI tools, nor was the analysis of the results obtained, either during the development of the code necessary for the machine learning analysis.

Once translated, a spelling and grammar check was performed using AI. No sources were obtained through AI.

In general, no process was based on AI tools, except for the style analysis, which was always verified in detail by the authors before accepting any corrections.

REFERENCES

Albiz, J., Viberg, O., & Matviienko, A. (2023). *Guiding visual attention on 2D screens: Effects of gaze cues from avatars and humans* [Conference article]. 2023 ACM Symposium on Spatial User Interaction. https://doi.org/10.1145/3607822.3614529

Ali, A., Bin Faheem, Z., Waseem, M., Draz, U., Safdar, Z., Hussain, S., & Yaseen, S. (2020). *Systematic review: A state of art ML based clustering algorithms for data mining* [Conference article]. 23rd IEEE International Multi-Topic Conference, INMIC 2020. https://doi.org/10.1109/INMIC50486.2020.9318060

Bednarik, R., & Tukiainen, M. (2006). *An eye-tracking methodology for characterizing program comprehension processes*.

Besl, P. J., & Jain, R. C. (1985). Three-dimensional object recognition. *ACM Computing Surveys, 17*(1), 430451. https://doi.org/10.1145/4078.4081

Bisley, J. W. (2011). The neural basis of visual attention. *The Journal of Physiology*, *589*(1), 49-57. https://doi.org/10.1113/JPHYSIOL.2010.192666

Bruce, C. D., & Hawes, Z. (2015). The role of 2D and 3D mental rotation in mathematics for young children: what is it? Why does it matter? And what can we do about it? *ZDM Mathematics Education*, *47*(3), 331-343. https://doi.org/10.1007/s11858-014-0637-4

Burgos, J. Manuel. (2017). *Historia de la psicología*. Palabra.

Calderón Ynoñan, P., Prieto Neira, F., & Sialer Rivera, M. (2022). *Integración de un enfoque ágil con técnicas de Diseño Centrado en Usuario (DCU) para la mejora de experiencia de usuario (UX)* [Undergraduate thesis, Universidad Señor de Sipán]. https://hdl.handle.net/20.500.12802/10085

Chelloug, S. A., Ashfaq, H., Alsuhibany, S. A., Shorfuzzaman, M., Alsufyani, A., Jalal, A., & Park, J. (2022). Real objects understanding using 3D haptic virtual reality for e-learning education. *Computers, Materials & Continua*, *74*(1), 1607-1624. https://doi.org/10.32604/CMC.2023.032245

Chochoiek, N. (2017). Explaining the success of user-centered design – An empirical study across German B2C firms. *Junior Management Science*, *2*(1), 81-116. https://doi.org/10.5282/jums/v2i1pp81-116

Correa Torres, Á. (2021). *Factores humanos y ergonomía cognitiva*. Universidad de Granada.

del Punta, J. A., Betegón, E., Gasaneo, G., Irurtia, M. J., & de Luis García, R. (2024). Contribuciones del rastreo ocular al test de percepción de diferencias: Un estudio en población universitaria. *Psykhe*, *33*(2), 1-18. https://doi.org/10.7764/PSYKHE.2022.46705

Duchowski, A. (2017). *Eye tracking methodology* (3rd ed.). Springer Nature. https://doi.org/10.1007/978-3-319-57883-5

Eke, C. I., Norman, A. A., Shuib, L., & Nweke, H. F. (2019). A survey of user profiling: State-of-the-art, challenges, and solutions. *IEEE Access*, *7*, 144907-144924. https://doi.org/10.1109/ACCESS.2019.2944243

Ferreres, A. (2023). *Cátedra I de neurofisiología*. https://www.psi.uba.ar/academica/carrerasdegrado/psicologia/sitios_catedras/obligatorias/048_neuro1/cursada/bibliografia.pdf

Gorichanaz, T. (2024). Identifying the barriers to human-centered design in the workplace: Perspectives from UX professionals. *Proceedings of the ACM on Human-Computer Interaction, 9*(2), CSCW110. https://doi.org/10.1145/3711008

Görücü, S., Morais, L. A., & Panagiotidou, G. (2025, April 26). *A critical analysis of machine learning eco-feedback tools through the lens of sustainable HCI* [Conference article]. Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3706598.3713198

Hernández Salazar, P. (1993). *El perfil del usuario de información*. https://ru.iibi.unam.mx/jspui/bitstream/IIBI_UNAM/A435/1/ibi000701502.pdf

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*(6), 787-795.

Jiménez, J., Colaboradores, S., Turvey, B. S., Ponce, A. C., Strano, M., Enrique, J., Castro, S., Canter, D. V, Youngs, D. E., Luis González Álvarez, J., & Sotoca Plaza, A. (2012). *Manual práctico del perfil criminológico: CRIMINAL PROFILING* (2a ed.). LEX NOVA.

Kashtalian, A., & Sochor, T. (2021). *K-means clustering of honeynet data with unsupervised representation learning* [Conference article]. IntelITSIS'2021: 2nd International Workshop on Intelligent Information Technologies and Systems of Information Security. https://ceur-ws.org/Vol-2853/paper48.pdf

Kumar, K., Chen, Y., Hu, B., & Luo, Y. (2024). *Assessing human visual attention in retail human robot interaction: A YOLOv8-Nano and eye-tracking approach* [Conference article]. 32nd ACM Conference on User Modeling, Adaptation and Personalization. https://doi.org/10.1145/3631700.3665195

Leivi, A. E. (2019). *Análisis de la implementación de Machine Learning en el diagnóstico por imágenes*. Universidad de San Andrés.

Liu, R., Menon, S., Mao, C., Park, D., Stent, S., & Vondrick, C. (2022). *Shadows shed light on 3D objects*. https://doi.org/10.48550/arXiv.2206.08990

Madi, N. AL. (2025). Identifying eye movement patterns for an adaptive approach to correcting eye tracking data in reading tasks. *Proceedings of the ACM on Human-Computer Interaction, 9*(3), ETRA13. https://doi.org/10.1145/3725838

Mondal, S., Pratim Das, P., & Bhattacharjee Rudra, T. (2022). Measuring code comprehension effort using code reading pattern. *Sādhanā, 47*, 117. https://doi.org/10.1007/s12046-022-01876-5

Nagro, S., & Aldekhail, M. (2021). Evaluation of multimedia user interface design method for m-learning (MobLearn): A comparative study. *Informatica*, *45*, 633-641. https://doi.org/10.31449/inf.v45i4.3391

Romanowski, R. (Ed.) (2019). *Managing economic innovations – Methods and instruments*. Bogucki Wydawnictwo Naukowe. https://doi.org/10.12657/9788379862771

Rueda Cuerva, C. (2021). *Educar la atención: con cerebro*. Alianza Editorial.

Sun, H., Zhu, F., Li, Y., Zhao, P., Kong, Y., Wang, J., Wan, Y., & Fu, S. (2023) . Viewpoint planning with transition management for active object recognition. *Frontiers in Neurobiotics, 17*, 1093132. https://doi.org/10.3389/fnbot.2023.1093132

Tarazona Evangelista, L., Aguilar Baca, Y., & Rivera Flores, V. (2018). *Uso de la teoría del color para mejorar la atención visual en estudiantes del segundo grado de secundaria de la I.E. San Martín de de Socabaya, Arequipa 2017* [Undergraduate thesis, Universidad Nacional San Agustín de Arequipa]. http://repositorio.unsa.edu.pe/handle/UNSA/6504

Téllez Rojas, M. A., Vila Carranza, N. J., Rojas Iturria, F. A., & Jiménez Castillo, R. (2021). Atención visual a las señales de consumo de tabaco: un estudio preliminar empleando el seguimiento ocular. *Revista Internacional de Investigación en Adicciones*, *7*(1), 16-24. https://doi.org/10.28931/riiad.2021.1.03

Thakur, A. (2020). *Approaching (almost) any machine learning problem*. https://github.com/abhishekkrthakur/approachingalmost

Ustárroz, J., & Grandi, F. (2016). Sobre la memoria de trabajo y la memoria declarativa: propuesta de una clarificación conceptual. *Cuadernos de NeuropsicologíaPanamerican Journal of Neuropsychology*, *10*, 13-31. https://doi.org/10.7714/CNPS/10.3.201

Uusitalo, S., Salovaara, A., Jokela, T., & Salmimaa, M. (2024). *"Clay to play with": Generative AI Tools in UX and industrial design practice* [Conference article]. ACM Designing Interactive Systems Conference, DIS 2024. https://doi.org/10.1145/3643834.3661624

Villegas Ortíz, A., Rodríguez López, E., & Álvarez Rodríguez, F. (2024). *Modelo para la aplicación del análisis del comportamiento de mirada mediante eyetracking de la interacción usuario-sistema para mejorar la calidad de la UX y mejorar los procesos de desarrollo de software* [Master's thesis, Universidad Autónoma de Aguascalientes]. http://hdl.handle.net/11317/3001

Wu, W., Ghazali, M., & Hazlin Huspi, S. (2024). A review of user profiling based on social networks. *IEEE Access, 12*, 122642-122670. https://doi.org/10.1109/ACCESS.2024.3430987