HELLENIC REPUBLIC
UNIVERSITY OF MACEDONIA

INTERDEPARTMENTAL PROGRAM OF POSTGRADUATE STUDIES IN INFORMATION SYSTEMS

MASTER THESIS

“PREDICTING CONSUMER EMOTIONS THROUGH MOUSE TRACKING ON GAMIFIED MARKETING CAMPAIGNS”

PASCHOU THEODORA

SUPERVISOR: PROTOGEROS NICOLAOS

Submitted as required to obtain the master’s degree
Diploma in Information Systems

OCTOBER 2022

ACKNOWLEDGEMENTS

I want to express my gratitude to everyone who helped me and supported me.

I must first express my gratitude to my research supervisor, Prof. Protogeros Nicolaos, whose knowledge and experience were crucial in developing the study questions and methods. Your vast expertise has been a huge help to me.

Additionally, I would like to convey my appreciation to Dr. Tzafilkou Aikaterini and Prof. Economides Anastasios for all their help and assistance with this research. During our discussions about this project, they brought up many insightful issues, and I hope I was able to answer a few of them here.

I wish to express my deepest gratitude to my family and friends whose loving support helped me to stay motivated and achieve my goal.

Last but not least, a big thank you to my friends for their wise counsel and my fellow students whose help and support were crucial for me throughout our studies.

Abstract

Over time, new practices have been developed as a result of the technological progress. Gamification is an excellent example. Gamification is a relatively new concept that has grown in popularity over time. Since it has been demonstrated that gamification may alter users' emotions, marketing is among the important industries that have embraced gamified tactics to enhance its aims.

In our study, mouse tracking technology was used to extract mouse interactions in gamified marketing campaigns. In particular, 132 individuals' cursor movements were implicitly monitored as they engaged on two gamified marketing tasks of a branded product. A set of mouse features was calculated, and two emotions (anger and frustration) were measured through self-reports. With the use of the mouse features, we could show how mouse tracking may help us forecast user emotions like anger and frustration. Four machine learning models were deployed. Logistic Regression, Random Forest, Logistic Regression, and XGboost were able to predict users' anger and frustration with accuracy scores of more than 80 %. The findings demonstrate that mouse tracking features can provide us predictions on users' emotions.

Our findings have relevant implications for a variety of domains where the capturing of mouse movements based on gamification and the predictions of users' sentiment and behavior could have a positive impact.

Table of Contents

Table Of figures

1 Introduction

This paper introduces an experimental attempt for predicting users' hesitation and confusion on a gamified marketing campaign, through mouse input, using Machine Learning algorithms. In this chapter, we will discuss the definitions of Mouse Tracking and Gamification, analyzing how they can be combined in the field of Marketing and Business.

More and more businesses in 2022 are starting to gamify different aspects of their work. While gamification has gained significant recognition and is applied in various fields, it can be described as a novel concept. This is because its definition was first mentioned back in 2002 by Nick Pelling (Christians 2018).

Therefore, there are several fields that have not used gamification in their research yet. For example, the application of mouse tracking in a gamification task is limited since there are not many works that have combined these two techniques and their technologies.What is more, gamification and mouse tracking have not been combined with machine learning as well. Thus, it has not been yet reported whether it is plausible to form predictions or to extract useful information combining gamification and mouse tracking in machine learning.

For above reasons, this research applied gamification techniques in the field of marketing instead of other common techniques, to emphasize and support the fact that the strategy of gamification has gained widespread acceptance and its use could provide many technological sectors with benefits and enhance their capabilities.

What is more, this study seeks to examine the extent to which, useful mouse features can be stored while users interact with a gamified task and whether these features are useful enough to be fed on machine learning algorithms and form predictions about users' emotions. Moreover, we discuss and explain the definitions of gamification and mouse tracking, reporting their importance in other fields especially in marketing and business.

1.1 Gamification Definition

In many research studies, Gamification is defined as the usage of design elements for games in non-game contexts, (Arnedo-Moreno et al., 2017; Høgenhaug 2012;

Triantoro et al., 2019). Others describe gamification as the process which combines game thinking and game elements to engage users and solve problems, in activities that are not games (Angelova, Kiryakova & Yordanova, 2015). According to Grif & Farcas (2016) gamified systems look like games, but they are not.

In his work, Xouridas (2020) proposes the definition of Adamou B. (2018) who states that "gamification borrows superficial aspects of games to win on the engagement that games can create for activities that are not games". Thus, gamification borrows game-based elements, and components to increase engagement and satisfy individuals' psychological needs, as games do. This tends to make gamification a feature that is inherently engaging the user, which means that it naturally captures attention and willingness, resulting in satisfaction and enjoyment for the user.

Gamification has raised a lot of interest in different fields, such as e-learning, education (Kalogiannakis, Papadakis & Zourmpakis, 2021; Hsin-Yuan Huang & Soman, 2013; Kim et al.,2018) business and marketing (Kaarlehto, 2020).

1.2 Gamification In business and Marketing

Gamification is a rapidly growing practice in the business world (Kaarlehto, 2020).

As reported by Gartner (2012), "gamification is the use of game mechanics to drive engagement in non-game business scenarios and to change behaviors in a target audience to achieve business outcomes."

To enhance the customer experience, businesses utilize gaming techniques and game-based rewards. For several years, industries, have used gamification systems to offer rewards to customers, such as earning points that can lead to a discount, free products, or exclusive offers. These programs are designed to increase audience engagement and loyalty by encouraging them to use or purchase the manufacturers' products and services to earn a reward (Giannakopoulou, 2020).

Businesses focus on marketing to develop new methods of promoting brands and helping them stand out. The significant part of these processes is that they are designed to provide the customer with a fun experience that includes the positive feelings provided by a game. As it is accepted, gamification is related to three main marketing goals: engagement, brand loyalty, and awareness (Lucassen & Jansen 2014). As a result, gamification can assist marketing in creating positive emotions in customers and achieve its basic goals.

Gamification uses techniques, which can alter audience's behavior in company business strategies, to enhance brand engagement (Gartner Research, 2011). According to Sailer et al., (2017) the primary goal of gamification, or the use of game design elements in non-gaming contexts, is to increase user's motivation and performance in each activity. Hence, the goal of gamification is to turn an event or a platform into game using game-based elements, to boost users' engagement and to enhance their experience.

To provide motivation to users and to examine the impact that gamification could have on users' engagement, some researchers like Xouridas (2020) included game design features on his experiment which was a gamified marketing campaign. For our experiment, we used the same technique gamifying our tasks.

1.3 Mouse tracking Definition and Applications

According to Wikipedia: "Mouse tracking (also known as cursor tracking) is the use of software to collect users' mouse cursor positions on the computer."

Generally, it has been proven that hand gestures and moves are influenced by emotions (Vicarion & Newman, 2013). As it is mentioned by (Arapakis, Lalmas, & Valkanas, 2014), "Usage of the mouse device can be thought of as consisting of a series of moves, aka gestures. Each such gesture is a specific and continuous physical process that is initiated and concluded by the user".

What is more, the use of the cursor is directly related to body activities and hand gestures. Consequently, in many research works it is suggested that cursor's activities are reflected from users' emotional states (Yamauchi, 2013).

In recent years, Mouse Tracking has been used in a variety of studies across different fields to gain insight and explain the link between cursor's activity and psychological aspects of users (behavior, emotions etc.). What is more, many predictions concerning users' emotional state have been formed, supported by mouse activity. In some cases (Banholzer et al., 2021; Yamauchi 2013), the feasibility of tracking mouse activity to measure stress and anxiety was explored. Moreover, many researchers applied mouse tracking to examine the link between cursor activity and user's behavior (Jaiswal,Tiwari& Hossain, 2020; Tzafilkou & Protogeros, 2018).

In other cases, mouse tracking has been applied to examine or diagnose user cognitive processes (Schoemann et al., 2020).

In order to measure and predict user engagement on web and different search engines, researchers have applied mouse tracking techniques to understand the way and the level users interact with web content (Arapakis et al., 2014 ; Arapakis & Leiva 2016; Konstan, J. A., Chi, E. H., Höök, K). A similar work comes Dias et al., (2013) who study the way that users interact with search engine result pages, while Chen et. al (2017) used cursor's activity to predict users' satisfaction.

On the same field, Tanjim-Al-akib et al. (2017) examined and described how websites can be improved, tracking and analyzing users' mouse activity. Mouse tracking has been also utilized to improve users' experience like in the case of Souza et al., (2019). Other researchers, have studied mouse behavior as a mean to detect or authenticate the user's trusted interaction behavior (Yi et al., 2020) or to elicit discussion about a concerning privacy issue relating to web browsing (Leiva, Arapakis & Iordanou, 2021)

Despite the broad application of mouse tracking in several scientific fields, there is not enough research in the context of interactive and gamified digital marketing campaigns.

Motivated by the above this study seeks to explore the link between users' emotion and mouse features on gamified environments. What is more, our basic aim was to discover if it is possible to predict two basic emotion, confusion, and anger using mouse features and applying machine learning algorithms.

The findings of this work seem promising and may provide gamification community with knowledge and motivation to study further the potential impact of mouse tracking and gamification.

2 Related Work

In recent years, the technique of mouse tracking has been spread and applied in numerous fields. Many researchers have used it to investigate and correlate users' emotions or behavior with cursor movements, whereas others have adopted its application to form predictions. A variety of experiments have been conducted in which users interacted with computer cursors while a wide range of data analysis techniques have been performed, the most common of which is statistics and machine learning.

An interesting study comes from Tzafilkou & Protogeros (2018). They examined the relationship between mouse patterns that are related to a user's behavior (number of mouse movements, number of clicks etc.) and a collection of EUD behavioral attributes (Self-Efficacy, Risk Perception, Willingness to Learn, Perceived Usefulness, Perceived Ease of Use). Applying descriptive statistics and Pearson correlation analysis, it turned out that there is a close relationship between mouse measurable attributes and EUD behavioral states.

Dias da Silva & Postma (2020) examined whether it is feasible to predict mind wandering with the utilization of mouse tracking. A memory operation span exercise was performed by 272 students as part of the experiment. Mouse features were reported for every participant. Then, Naive Bayes, Linear Discriminant Analyses, K-Nearest Neighbors, Tree Bag, and Random Forest classifiers were used to predict mind wandering. This study established that hand gestures can indicate mind wandering.

Using the impact of information foraging theory, Jaiswal et al. (2020) achieved to predict users' behavior on web by tracking their mouse activity. Recording the activity of ten users who performed unknown search tasks, they tracked the position and time-based information of the mouse pointer. Specifically, a model was created, using RNNs (Recurrent neural networks) to identify the main correlations of the user's movement on the site and to analyze the behavior of the user's mouse movement on any website. Then, users' behavior was effectively predicted using Long short-term memory (LSTM) RNNs model.

Yamauchi (2013); Yamauchi & Xiao (2018), investigated whether mouse trajectories features can predict users' state anxiety. In this research, 234 people took part in the first experiment, which was the feature selection part. During the analysis, 134 mouse features were extracted by the Boruta algorithm. Secondly, 133 users participated in the evaluation experiment. Both experiments were choice-select tasks. Features from the primary task are used to feed the support vector regression. Results showed that the extracted features could predict users' anxiety scores successfully.

According to Yamauchi and Xiao (2018), mouse motions can provide information about users' emotions. For this research, four experiments were performed and completed. The first experiment was the correlation study between users' levels of anxiety and mouse trajectories. Next, three more experiments are proposed to examine the link between emotions and mouse motions further. Specifically, a music-based, film-based, and picture-based experiment took place to elicit users' emotions. Random Forest and Support Vector Machine were used to measure the extent to which mouse trajectories extracted from the first experiment, could predict elicited emotions.

Finally, the above study revealed that the extracted mouse trajectories can provide knowledge about users' emotions and predict the emotions of new users.

Banholzer et al. (2021) tried to prove that mouse movements are linked to workplace stress. In a prominent European technology company, all the computer mouse movements of seventy-one employees were recorded for about 30 minutes every day. Participants were then asked to rate their degree of stress. The mouse features that stood out were mouse speed and accuracy. To examine if work stress is linked to a speed-accuracy trade-off in computer mouse movements, the researchers used a

Bayesian regression model. Finally, the model revealed that stress had a negative relationship with mouse speed and accuracy. As a result, this analysis proved that work stress is associated with a speed-accuracy trade-off.

In e-learning applications of mouse tracking, as in the case of Tzafilkou & Protogeros (2020), cursor's activity was monitored and linked to students' acceptance items of perceived usefulness and ease of use. Solving a learning task, thirty students participated in the research. Mouse metrics that were captured were the number of mouse clicks, number of mouse hovers, number of mouse hovers that turned into clicks, and duration of mouse hovers before mouse clicks. Following that, their acceptance items (PU and EOU) were gathered from their post-experiment questionnaire responses. Descriptive analytics and Pearson correlation revealed that there is a significant link between mouse metrics and acceptance items.

In the same field, Rodrigues et al. (2013) found a relationship between the stress level of students and the number of mouse usage. For research shake, ten students completed the same assignment with and without restrictions. Results showed that when pupils were stressed, there was a considerably higher number of mouse movements. Hence, there is a strong relationship between students' stress levels and their mouse movements.

Mouse-tracking recording techniques are also gaining popularity in (UX). Tzafilkou et al. (2014) used mouse tracking to explore the effects of two behavioral variables, self-efficacy, and risk perception, on user experience and user performance. Thirty-two end users created their own applications using a web-based tool. Next, they answer questions related to their perceived risk and self-efficacy. They evaluated mouse motions as behavioral findings in order to investigate the association between cursor movements and user behavior. Specifically, apart from mouse movements, the hesitation level was tracked as well. Then mouse hesitation and questionnaire hesitation were compared, concluding that there is a close relationship between them. Therefore, this study demonstrates that mouse tracking merged with behavioral analysis can reveal useful information about a user's experience.

Aviz et al. (2019) compares mouse and gaze tracking approaches in a comparative study of UX evaluation techniques. In the experiment, 10 computer engineers participated, completing 4 tasks on an economic tax-related website of the country.

recording the task completion time, where two types of users' behavior were observed, dynamic and statistical (those who interact with the cursor dynamic and those who leave it static). Next, the Euclidean distance between the gaze and the mouse was also captured and analyzed. The above observations, in combination with eye heat maps, revealed that users with dynamic browsing behaviors tend to have the same direction of gaze and mouse, and hence, their task execution time is less than users who engage in static browsing behavior with a fixed mouse.

Although mouse tracking is a technique that is constantly evolving and gaining great publicity, it has not been combined in many cases with gamification. One of the few surveys that combined gamification and mouse tracking is the case of Betz et al. (2020). They introduced a novel approach to mouse-tracking for analyzing online voice processing. They challenged existing paradigms by including the work in a drag-and-drop game that provides performance feedback in the form of a score. The primary purpose of this paper is to investigate the potential of Mouse Tracking when it is enhanced with gamification elements. Their findings are only useful for deducing initial tendencies that could be used as hypotheses for further research, but they are sufficient as proof of concept that their gamified mouse-tracking system is suitable for studying online speech processing. The experiment has been designed like a game. While stimuli were given, 8 participants were taught to navigate as precisely and swiftly as possible between abstract and concrete items. The three audio circumstances were: NO (no hesitation), LEN (hesitation lengthening), and FULL (lengthening and filler). Finally, it was discovered that the condition of hesitation was greater in abstract targets than in real targets. The contrary was true in real targets. An important finding was that, in general, the Mouse tracking and gamification approaches function satisfactorily in terms of providing meaningful and analyzable outcomes.

In conclusion, these studies all suggest an important link between a user's behavior and mouse activities. However, there are still several critical problems to be clarified. First, the data relating mouse behavior to emotion is still shaky. Second, most studies employ a small number of participants and small data sets. Thus, the statistical power of these studies is negligible. Third, few studies have combined mouse tracking technology with gamification. Finally, the general link between users' feelings and gamification, applying machine learning algorithms, has not been investigated in many cases.

With these issues in mind, the focus of the current study is to predict users' anger and confusion. Using a satisfactory number of participants who interacted with two different games, we managed to collect mouse activities which used as features in the Logistic Regression and Random Forest algorithms.

Table 1 Extracted mouse metrics/features Target Variables and Number of Participants

Authors and Year	**Mouse Metrics/Features**	**Target Variable**	**Number of Participants**	**Method**
Banholzer et al. (2021)	Mouse Speed Mouse Accuracy Mouse Clicks Mouse Wheels	Work Stress	71	Correlation
Dias Da Silva & Postma (2020)	x-pos max, x-pos min, y-pos max, y-pos min, MAD, MAD time, MD above, MD above time, MD below, MD below time, AD, AUCb x 103, x-pos flips, y-pos flips, x-pos reversals, y-pos reversals, RT, initiation time, idle time, total dist, vel max, vel max time, acc max, acc max time, acc min, acc min time, sample entropy, set size	Mind Wondering	272	Prediction
Tzafilkou & Protogeros (2020)	Number of mouse clicks, Number of mouse hovers, Number of mouse hovers that turned into clicks, Duration of mouse hovers before mouse clicks	Perceived ease use and usefulness	30	Correlation Analysis
Jaiswal et al. (2020)	users' click features in X– Y direction in all sessions, state-wise distribution of mouse in X–Y direction associated with users' mouse positions and click events, distribution of user-wise distinctive X– Y movements	User mouse movement patterns	10	Prediction
Yi et al. (2020)	Type of the mouse event, x-coordinate of the mouse pointer (x),	The accuracy of the user's trusted	8 18	Prediction

	y-co-ordinate of the mouse pointer (y), time of the mouse event (t)	interaction behavior identification		
Aviz et al. (2019)	Execution time of each task, Average distance between eye gaze focus and mouse cursor, mouse cursor stops, heatmap; (c) mouse cursor path followed.	Type of behavior	10	Correlation
Tzafilkou & Protogeros (2018)	Number of mouse movements in a session, Number of non-direct (curved) movements in a session, Number of clicks outside direct movements (lines) in a session, Number of clicks in the end of direct movements (lines) in a session, Number of mouse long pauses in a session, Ratio of average time of long pauses to the task duration time, Number of direct movements in a session, Ratio of the average pauses time between direct movements to the task duration time, Number of mouse hovers that turned into mouse clicks in a session, Average time from mouse hover to mouse click on the same element, Ratio of average time between clicks to the task duration time, Ratio of average time of pauses to the task duration time, Ratio of average time of pauses direct movements in a session, Number of slow movements in a session, Number of all mouse hovers in a session, The EUD task total duration (completion) time (which is different for every user),Activity level for mouse movements	Self-Efficacy (SE) Risk Perception (RP) Willingness To Learn (WL) Perceived Usefulness (PU) Perceived Ease of Use (PEOU)	30	Correlation
Tzafilou et al. (2014)	Mouse movements, Mouse Clicks	Hesitation	32	Correlation
Rodrigues et al. (2013)	Mouse down, Mouse up, mouse wheel, Number of mouse movement	Stress Level	10	Correlation
Yamauchi (2013)	For female subjects: velocity (skewness), direction change, velocity (kurtosis), velocity (sd), velocity (kurtosis), attraction+, x-overshoot, path length-x, attraction+ For male subjects: mid-line cross, decision speed, end time, velocity	State anxiety score	234 133	Prediction

(mean), path length-y, direction change, decision speed, velocity (kurtosis), velocity (skewness), mid-line cross

Table 2 Algorithms and Scores

Authors and Year	**Algorithms**	**Results/Scores**
Dias Da Silva & Postma (2020)	K-nearest Neighbor (KNN) Random Forest (RF) Linear Discriminant Analysis Naive Bayes Tree Bag Random Forest	Acc=0.38 Acc=0.40 Acc=0.37 Acc=0.47 Acc=0.42 Acc=0.47
Jaiswal et al. (2020)	Recurrent neural networks (RNNs) Long short-term memory RNNs (LSTM)	mean squared error (MSE)=0.157 root mean squared error (RMSE)=0.396
Yi et al.	Random Forest	Accuracy=91.82% Error rate<8.18%
Yamauchi (2013)	Support Vector Regression	0.28<=Prediction Performance Score<=0.63

3 Methodology

3.1 Participants and Procedure

For the experiment, about 132 participants were gathered. Due to the COVID pandemic and since at that time Greece was in quarantine, each participant completed the experiment remotely, using their individual computer. 96 participants were students enrolled in the postgraduate program in information systems at the University of Macedonia (Greece), and in the postgraduate program at the International Hellenic University (Greece). The other 36 people were random participants, invited through email messages from the authors. During their online university course, students have been instructed to complete the gamified task. The rest of the group completed it at a random time. They were all native Greek speakers with good computer knowledge. There were no notable gender or age contrasts between participants. About 21 participants were excluded because they did not complete the task properly. All participants received an instruction page that could be accessed online through a URL. To successfully complete the task, they needed to log in to the EasyPromos Platform (the credentials were provided in the instructions), play the first game, respond to a self-reported questionnaire, and then play the second game and respond to the second self-reported questionnaire. Each game lasted one minute, and the whole procedure did not last more than five minutes.

The self-reported questionnaire was embedded in the game as an online form, where the users had to reply to a set of questions about their anger and confusion rates. Specifically, a five-point Likert-type scale was used to measure participants' anger and confusion levels.

The experiment was being conducted for almost one month. Every user needed to complete 2 gamified tasks and reply to the corresponding questionnaire. Although some of them completed only one of the two tasks, we kept their data. The age range was varied, from 21 to 30 years old. There was no noteworthy gender difference between participants. All data were processed anonymously, and all participants provided their informed consent before participating in the study.

3.2 The experiment

For the current research, an experiment was carried out, which has been designed as a marketing campaign, to provide a more realistic perspective. Specifically, participants were to win a prize once they completed the two games within a certain time. For the campaign, we chose a well-known soft drink company because we assumed that its logo was one of the most recognizable in the world and, thus, it could attract more attention from users.

The experiment consisted of two gamified tasks. The first one was a hidden items game and the second was a puzzle game. The first gamified task was problematic, and it did not have a solution. Specifically, while participants were asked to detect five bottles of a known soft drink, the game was designed so that the player could not step on a specific bottle object. The result was that one pressed on the object many times without being able to select it. Next figure depicts the Hidden-items game. The green areas in Figure 1 are the ones that could be clicked. At the top of the image, it appears in red, the bottle which could not be selected. This happened to evoke extreme feelings among the players.

The second game included an easy solution, and its solution was very simple. The difficulty of the two games varied enormously. Our intention was to stimulate two completely different levels of confusion and anger.

In the end, the self-reported questionnaire embedded in the game as an online form, where the users had to reply to a set of questions about their anger and confusion rates. Specifically, a five-point Likert-type scale was used to measure participants' anger and confusion levels.

The procedure was being conducted for almost one month. Every user needed to complete 2 gamified tasks and reply to the corresponding questionnaire. Although some of them completed only one of the two tasks, their data were kept as well. The age range was varied, from 21 to 30 years old. There was no noteworthy gender and ethnicity difference between participants thus we did not take them into account. All data were processed anonymously, and all participants provided their informed consent before participating in the study

Figure 1 Screenshot of the frustration inducing task

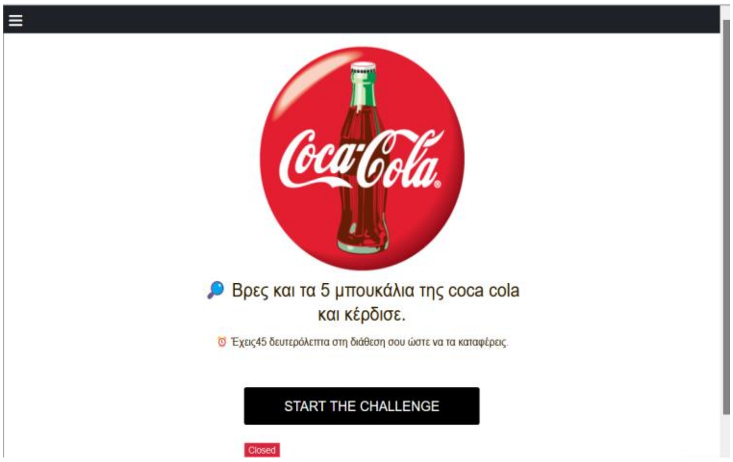

Figure 2 Screenshot of the Hidden items game start page

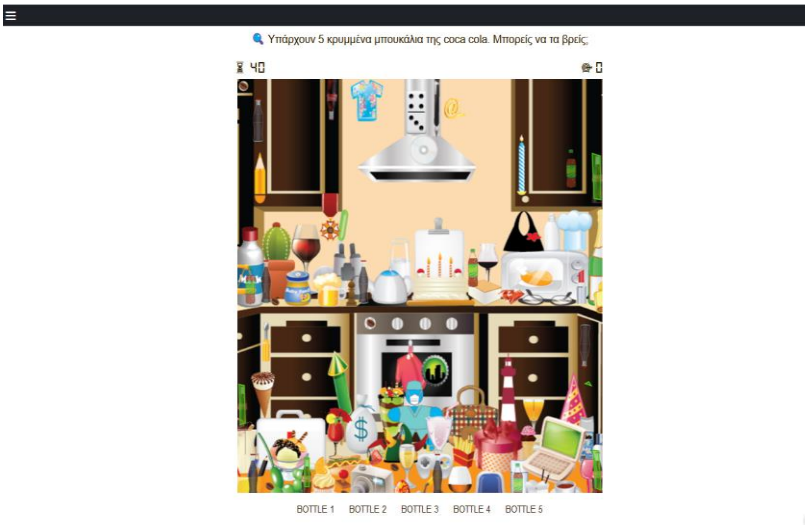

Figure 3 Screenshot of the Hidden items game on Easy Promo

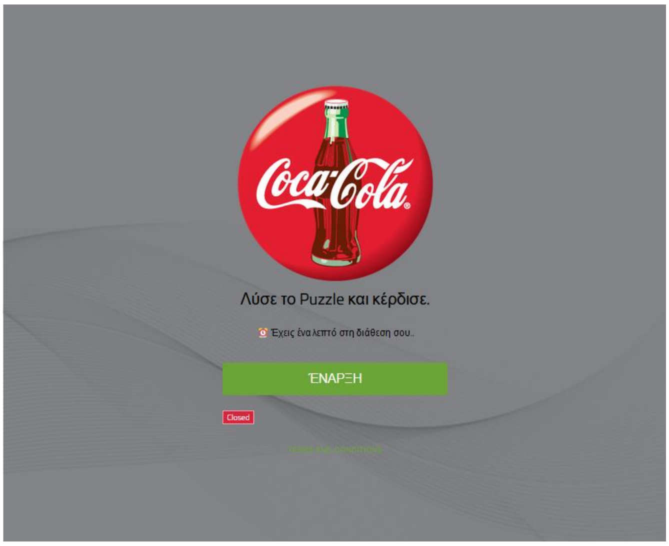

Figure 4 Screenshot of the Puzzle game start page

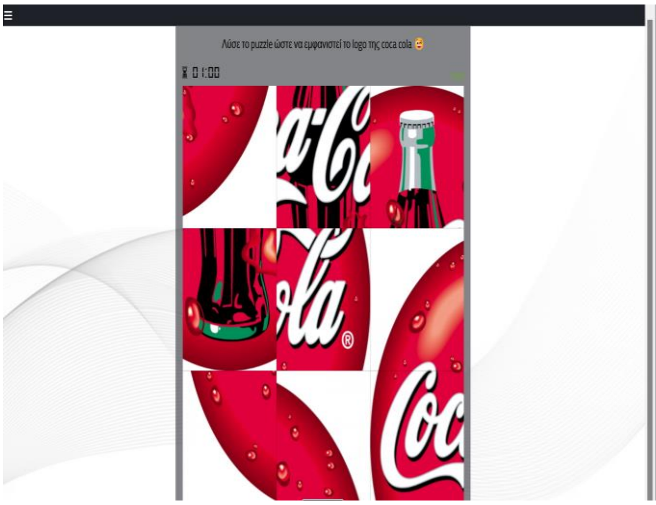

Figure 5 Screenshot of the Puzzle game on Easy Promo

3.3 Material and Technology

For the creation and designing of our gamified tasks, we used the Easypromos platform. Easypromos is an application for creating digital giveaways, surveys, contests, promotions, and games. As for the gamified solutions, it supports eight different types of games (Puzzle, Timed Quiz, Memory, Match it, Word Search, Slide and Match, Hidden Objects, Minesweeper). As regards the design of the Hidden-items game, we also used a free online graphical tool, Canva, to create the final image of the hidden items. First, we downloaded a free image from Google and then we edited it using Canva.com .

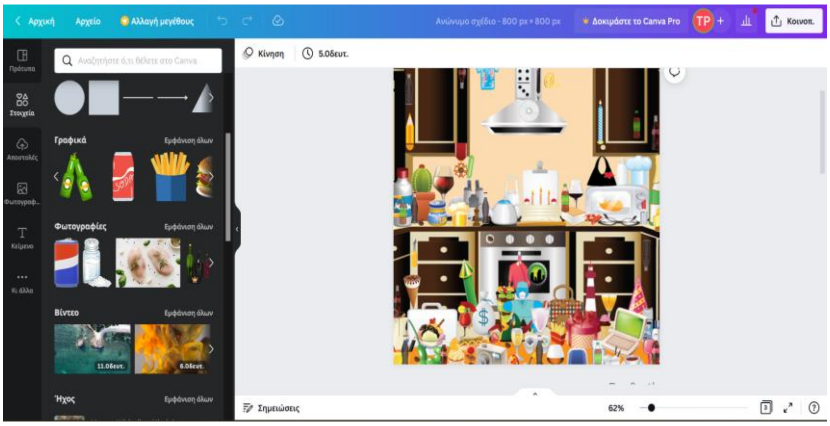

Figure 6 Canva Application

Easypromos is also a self-service platform for creating and managing digital promotions across any social network, web, or device, with 30 promotion apps available, including games, quizzes, social media giveaways, photo- and video contests, surveys, coupon codes, and more. Commonly, it is used for Social Media Marketing, Brand Management, Campaign Management, and Contests. It provides a wide variety of solutions, such as product promotion, customer loyalty, data collection, and gamification. Easypromos allows integration with platforms such as Salesforce, Facebook, Mailchip, Wordpress etc.

Since EasyPromos allows the integration with Wordpress and an easy presentation of promotions on websites or blog pages, we designed a Wordpress page in the form of Blog and connected it with the Gamified campaign.

Easypromos allow users to easily introduce contests, giveaways, games, and other promotional applications to any page of their choice, be it a blog post, sidebar, etc. In our case, the gamified campaigns were embedded through an html iframe in the content of a WordPress page. In the end, participants received the gamified tasks via URL. Then, they entered the WordPress environment to start the procedure.

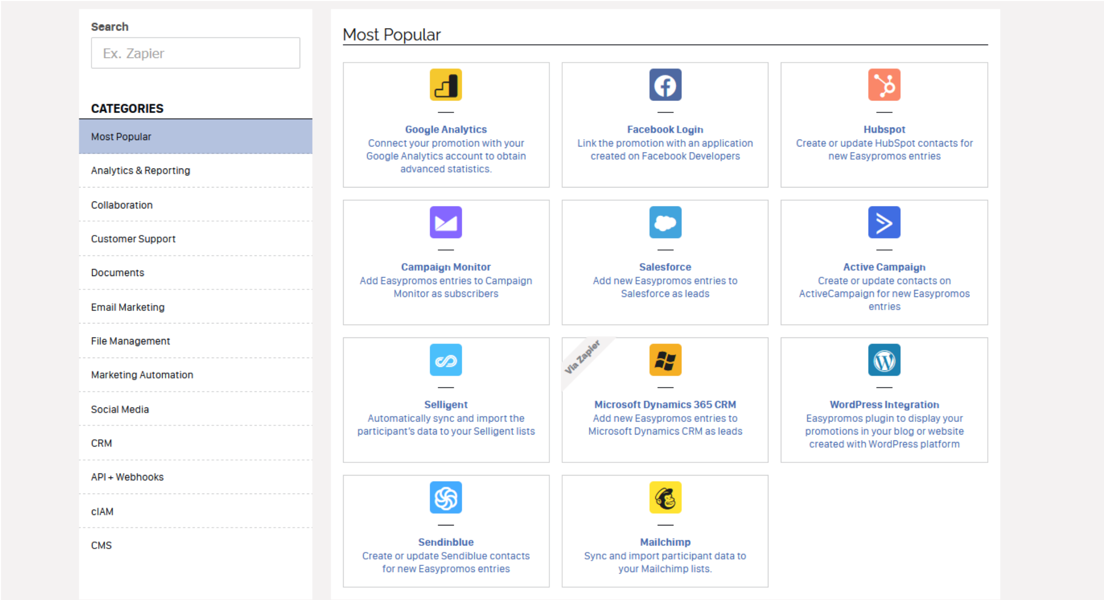

Figure 7 EasyPromos Integration

3.4 Mouse Tracking Mechanism and Mouse Data

A JavaScript-based prototype monitoring tool was developed to capture user mouse behavior during the user-web interaction and store in real time the captured events. Moreover, the Mouse Tracking mechanism is based on a JavaScript jQuery program that records time and space-based mouse events (click, move, etc.) and saves them in JSON format on a remote server. Saving is done with the ajax post method at the push of a button. Until the button is pressed, the information is stored locally in the local store of the browser.

The stored raw mouse data consisted of time-based and space-based mouse information. Specifically, the data that was tracked was information about Clicks (CoordX, CoordY, TimeStamp,TimeSince), Speed (Speed, Acceleration) and Moves (CoordX, CoordY, TimeSince). Additional information was also collected from the questionnaire (Angry, Confused, Game, Name). Due to a technical dispute, clicks were not recorded for every player. Thus, mouse click information was deleted in every JSON file, and only mouse movement data was saved.

3.5 Data Analysis and Mouse Features

For the mouse data processing, we first converted every JSON file to excel format. Thus, several free online conversion programs were used. Files with missing information were deleted. We kept only the JSON files of users who completed the entire game process. On average, there were two excel files for each player. While some users completed only one of two games correctly, others played a game more than once. However, their information was kept as well. Then, every excel file was uploaded to Jupiter notebook for the mouse features extraction process.

At the beginning of the process, Pandas package was imported. Pandas, is a Python package for data analysis and manipulation and is open source.

Pandas make it simple to create, manipulate, and display data in a data frame (Johnson, 2022).

Then, the excel file was red using .read_excel () method.

```
import pandas as pd
data = pd.read_excel('alexaki13_24_04_2021_14_37_15.xlsx')
data.describe
```

Figure 8 Importing Pandas package and reading the excel file

What is more, the variance of speed was calculated using the .var() method.

```python
var_speed= data['speed__Speed'].var()
print(var_speed)
```

```
150707.55858880107
```

Figure 9 Variance Calculation

Next, the number of pauses were calculated as well. If the time since next clicks was greater than 3000 seconds, it was considered as a pause.

```python
In [8]: pauses= data.loc[data["clicks__TimeSince"] > 3000]
        pauses.head(20)
```

Out[8]:

	clicks__CoordX	clicks__CoordY	clicks__Timestamp	speed__Speed	speed__Acceleration	moves__CoordX	moves__CoordY	ungry	confused	name	mouse
3	464.0	502.0	1.619275e+12	50.000000	-306.225775	740	185	NaN	NaN	0	NaN
4	574.0	175.0	1.619275e+12	0.000000	-500.000000	740	184	NaN	NaN	0	NaN
6	566.0	189.0	1.619275e+12	0.000000	0.000000	740	179	NaN	NaN	0	NaN
7	469.0	377.0	1.619275e+12	101.980390	1019.803903	740	178	NaN	NaN	0	NaN
8	566.0	171.0	1.619275e+12	98.994949	-29.854409	740	177	NaN	NaN	0	NaN

Figure 10 Calculation of the Number of Pauses

Most metrics were calculated using .describe() function, which calculates and present the summary of descriptive statistics for a data set.

```python
In [2]: data.describe()
```

Out[2]:

	clicks__CoordX	clicks__CoordY	clicks__Timestamp	speed__Speed	speed__Acceleration	moves__CoordX	moves__CoordY	ungry	confused	clicks__Tin
count	10.00000	10.000000	1.000000e+01	558.000000	557.000000	1525.000000	1525.000000	1.0	1.0	9
mean	621.50000	358.900000	1.619275e+12	169.939320	-0.538600	676.099016	368.446557	3.0	3.0	3728
std	154.38642	182.634699	1.223114e+04	388.210714	4652.945359	202.774628	163.307547	NaN	NaN	2873
min	464.00000	171.000000	1.619275e+12	0.000000	-32893.725620	351.000000	4.000000	3.0	3.0	952
25%	545.00000	179.250000	1.619275e+12	0.000000	-432.455532	530.000000	228.000000	3.0	3.0	1656
50%	566.00000	325.000000	1.619275e+12	30.000000	0.000000	625.000000	379.000000	3.0	3.0	3391
75%	664.75000	553.750000	1.619275e+12	175.188545	300.000000	784.000000	533.000000	3.0	3.0	4703
max	890.00000	580.000000	1.619275e+12	3556.205843	35162.058433	1267.000000	592.000000	3.0	3.0	10247

Figure 11 Descriptive analytics

Table 3 presents the final mouse metrics calculated and analyzed for each player. There were three more mouse metrics related to clicks (total number of clicks, number of pauses and mean time since next click). These features were deleted since they were not recorded for every player.

The last two mouse metrics, Confused and Angry, were recorded from the questionnaire which was appeared at the end of each game. The values they got were numbers from 0 to 5. However, we converted the values 0,1,2 and 3 to “no”, while the answers 4 and 5 got the value “yes”.

For every user, the calculated mouse metrics were stored in the final csv file. Each row represents a user while each column contains user’s mouse features. The final csv file contained 111 rows and 9 columns (Fig. 12)

Table 3 Extracted and calculated mouse metrics

Name of Metric	Description
Mean Speed	Average Speed of the mouse cursor during a user session
Mean Acceleration	Average Acceleration of cursor movements during a user session
Std Speed	Standard deviation of the mouse speed during a user session
Var Speed	Variance of the mouse speed during a user session.
Mean moves Times Since	Average of all times (in ms) between all mouse movements
Angry	Yes if the player stated anger at level 4 or more [1,5] No if the player stated anger at level 3 or less [1,5]
Confused	Yes if the player stated confusion at level 4 or more [1,5]

	No if the player stated confusion at level 3 or less [1,5]
Game	The type of Game (Hidden Items or Puzzle)

	A	B	C	D	E	F	G	H	I
1	player	mean_speed	mean_acceleration	std_speed	var_speed	mean_moves_time_since	ungry	confused	game
2	test	300.895.908	-2,19E-08	554.190.834	3.071.274.808.718.220	36.698421	no	no	PUZZLE
3	alexaki13	169.939320	-0.538600	388.210714	150707.55858880107	36.508780	no	no	hidden items
4	asteris	192.827256	-40.530885	453.341490	205518.506968307	45.695513	no	no	PUZZLE
5	asteris	158.568358	-33.899078	493.920591	243957.5498151956	52.820256	yes	yes	hidden items
6	athanasia1995	299.964914	-1.313469e-13	699.966911	489953.6769847257	53.182448	no	no	PUZZLE
7	athanasia1995	274.628967	-54.114887	661.814047	437997.8323682635	57.858696	yes	yes	hidden items
8	bagia	355.071723	-35.273815	654.828057	428799.78461432987	27.441459	no	no	PUZZLE
9	bagia	187.319038	-19.529569	447.631046	200373.5537450753	43.095814	yes	yes	hidden items
10	Danai123	422.704195	-74.463977	910.546740	829095.3665761523	47.957589	no	no	PUZZLE
11	Danai123	259.249124	-59.579498	708.629536	502155.8199144427	64.812233	yes	yes	hidden items
12	kostaskzn	271.221618	-49.193743	641.779954	411881.5092317408	69.535461	no	no	PUZZLE
13	kostaskzn	254.070318	-3.809524	686.245345	470932.67409836996	71.340031	no	no	PUZZLE
14	monika	331.408295	-128.910798	706.242314	498778.20599546237	46.619490	no	no	PUZZLE
15	monika	248.461652	-2.351925	580.155141	336579.9879282172	53.110285	yes	yes	hidden items
16	nikiii	534.456187	-39.054194	881.715021	777421.3788577815	23.891984	no	no	PUZZLE
17	nikiii	565.982160	-35.714286	1115.640845	1244654.49397045	47.111111	yes	yes	hidden items
18	nikosgiltidis	254.189609	-2.821537	606.949944	368388.23437143024	51.892442	no	no	PUZZLE
19	nikosgiltidis	128.597467	-9.447459	390.025827	152120.14539814196	46.612265	yes	yes	hidden items
20	tud123	271.011306	-93.359500	615.546443	378897.423121832	55.549724	no	no	PUZZLE
21	tud123	256.310783	-16.269173	614.543071	377663.1864433859	61.851604	yes	yes	hidden items
22	alexia	312.717604	-90.895050	624.421935	389902.75305195886	33.49796	no	no	PUZZLE
23	alex	157.992791	-96.399520	591.652740	350052.96446855523	104.644627	yes	yes	hidden items
24	theop13	393.927877	-185.538701	961.022679	923564.5900165279	50.749304	no	no	PUZZLE
25	thanasiskeramas	94.172582	-8.625258	281.575836	79284.95134731046	71.125951	no	no	PUZZLE
26	thanasiskeramas	233.066709	-26.309698	373.262833	139325.14226291914	32.744395	no	no	PUZZLE
27	stoik97	448.223313	-12.114280	880.130248	774629.2537722151	15.00795	no	no	hidden items

Figure 12 A sample of the csv file

3.6 Data Manipulation/Data Preparation

For the final data-preparation end the creation of Machine Learning prediction models, Kaggle.com was used. According to Wikipedia, “Kaggle is an online community of data scientists and machine learning practitioners. Kaggle allows users to find and publish data sets, explore, and build models in a web-based data-science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges”.

Since we wanted to predict both anger and confusion, two different notebooks were created. However, the process followed was the same for both of our target variables.

Below will be analyzed the whole procedure followed to bring the dataset into a suitable format, that can be used in the final predictive models.

3.6.1 The procedure

After storing the data in the csv file, the file was uploaded and read to Kaggle using the .read_csv() method.

```
#loading data
df=pd.read_csv("../input/mydata/kaggle.csv")
```

Figure 13 Data reading

Before that, all the necessary python packages were imported.

```
In [1]:
import numpy as np
import pandas as pd

import matplotlib.pyplot as plt
import seaborn as sns
import pandas_profiling as pp

from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report,accuracy_score,f1_score,precision_score,recall_score,roc_curve,roc_auc_score

import warnings
warnings.filterwarnings('ignore')
```

Figure 14 Python Packages importing

Our data set was consisted of nine columns: player, mean_speed, mean_acceleration, std_speed, var_speed, mean_moves_time_since, ungry, confused, game

```
df.columns
```

```
Index(['player', 'mean_speed', 'mean_acceleration', 'std_speed', 'var_speed',
       'mean_moves_time_since', 'ungry', 'confused', 'game '],
      dtype='object')
```

Figure 15 Dataset's columns

The data were composed of both numerical and categorical features.

- Numerical: mean_speed, mean_acceleration, std_speed, var_speed, mean_moves_time_since
- Categorical: angry, confused, game

```
mean_speed                float64
mean_acceleration         float64
std_speed                 float64
var_speed                 float64
mean_moves_time_since     float64
ungry                      object
confused                   object
game                       object
dtype: object
```

Figure 16 Data types

Next method, df.info(), prints information about a DataFrame including the index, the dtype, columns' name , non-null values and memory usage.

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 110 entries, 0 to 109
Data columns (total 9 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   player                 110 non-null    object
 1   mean_speed             110 non-null    object
 2   mean_acceleration      110 non-null    object
 3   std_speed              110 non-null    object
 4   var_speed              110 non-null    object
 5   mean_moves_time_since  110 non-null    float64
 6   ungry                  110 non-null    object
 7   confused               110 non-null    object
 8   game                   110 non-null    object
dtypes: float64(1), object(8)
memory usage: 7.9+ KB
```

Figure 17 Data information

The first column, included the user' name, was deleted since it was not useful input for our prediction models. Next figure presents the first 5 rows of our data set after deleting the first unnecessary column.

```
#delete users' name column
df.drop(df.columns[[0]], axis=1, inplace=True)
df.head()
```

	mean_speed	mean_acceleration	std_speed	var_speed	mean_moves_time_since	angry	confused	game
0	300.895.908	-2,19E-08	554.190.834	3.071.274.808.718.220	36.698421	no	no	PUZZLE
1	169.939320	-0.538600	388.210714	150707.55858880107	36.508780	no	no	hidden items
2	192.827256	-40.530885	453.341490	205518.506968307	45.695513	no	no	PUZZLE
3	158.568358	-33.899078	493.920591	243957.5498151956	52.820256	yes	yes	hidden items
4	299.964914	-1.313469e-13	699.966911	489953.6769847257	53.182448	no	no	PUZZLE

Figure 18 First five rows of our data set after deleting the first column

Angry and confused were the two target variables. Since we run our models twice, each time the target variable was the one of the two variables, depending on which emotion we wanted to predict.Next, mean_speed,mean_acceleraion, std_speed and var_speed data types were turned into numerical.

```
df.mean_speed = pd.to_numeric(df.mean_speed, errors='coerce')
```

```
df.mean_acceleration = pd.to_numeric(df.mean_acceleration, errors='coerce')
df.std_speed = pd.to_numeric(df.std_speed, errors='coerce')
df.var_speed = pd.to_numeric(df.var_speed, errors='coerce')
```

Figure 19 Converting data types

A very importing part in data preparation manipulation is searching for missing values and delete them.

```
df.isnull().sum()
```

Figure 20 Searching for missing values

```
df.dropna(inplace = True)
```

Figure 21 Deleting missing values

In the next step, the target variable was converted into 0 and 1.

```
df['confused'].replace(to_replace='yes', value=1, inplace=True)
df['confused'].replace(to_replace='no',  value=0, inplace=True)
```

Figure 22 Converting target variable into 0 and 1

What is more, categorical variables were converted into dummy variables

```
df_dummies = pd.get_dummies(df)
df_dummies.head()
```

Figure 23 Converting categorical variables into dummies

Figure 24 depicts the first 5 rows of our data set after the conversion of categorical variables into dummies and is the final dataset.

	mean_speed	mean_acceleration	std_speed	var_speed	mean_moves_time_since	confused	ungry_no	ungry_yes	game _PUZZLE	game _hidden items
1	169.939320	-5.386000e-01	388.210714	150707.558589	36.508780	0	1	0	0	1
2	192.827256	-4.053088e+01	453.341490	205518.506968	45.695513	0	1	0	1	0
3	158.568358	-3.389908e+01	493.920591	243957.549815	52.820256	1	0	1	0	1
4	299.964914	-1.313469e-13	699.966911	489953.676985	53.182448	0	1	0	1	0
5	274.628967	-5.411489e+01	661.814047	437997.832368	57.858696	1	0	1	0	1

Figure 24 Final data set

4 Modeling

In this section we analyze and describe 4 Machine Algorithms used in our research. First, the most common approaches are summarily described, pointing out the difference between Classification and Regression algorithms. Then, a brief analysis of our 4 Machine Algorithms is followed, which concludes the process of modeling and calculated predicting scores.

4.1 Machine Learning Algorithms

The type of Machine Learning can be categorized based on how the algorithm learns to improve its prediction accuracy. What is more Ayodele (2010), supports that based on the desired outcome of the algorithm, machine learning algorithms are classified into the next 4 categories.

- **Supervised machine learning:** This approach is defined by its use of labeled datasets. The model can be measured to improve its accuracy over time, using labeled inputs and outputs. These data set can be used to train or “supervise” algorithms in classifying data or predicting outcomes.
- **Unsupervised machine learning:** Use algorithms to analyze and cluster unlabeled data sets, allowing hidden patterns in data to be detected. The is no human supervision.
- **Semi-supervised learning:** in this approach both labeled and unlabeled examples are combined to create a classifier or function.
- **Reinforcement learning:** Based on a world observation, the algorithm is taught how to act. Every action impacts the environment, and the environment gives feedback to the learning algorithm.

The most common approaches are supervised and unsupervised. In Supervised machine learning, algorithms are divided into Classification and Regression Algorithms.

The main distinction between Classification and Regression algorithms is that Regression algorithms are used to determine continuous values such as price, income, etc. On the other hand, Classification algorithms are used to predict or classify distinct values such as True or False. Our problem ended up being a classification problem since we converted to “yes”, those answers that had a value of three and more while the rest got the value “no”.

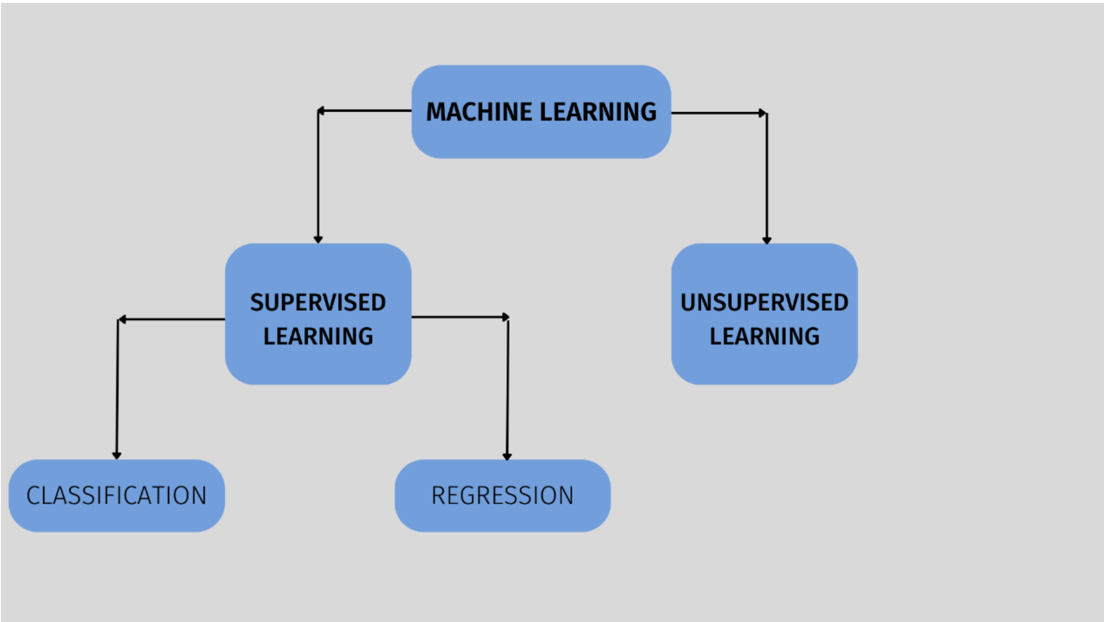

Figure 25 Supervised and Unsupervised Machine Learning approaches

According to Ayodele (2010), some examples of supervised machine learning algorithms that are more concerned with classification are: Linear Classifiers, Logistic Regression, Naïve Bayes Classifier, Perceptron, Support Vector Machine; Quadratic Classifiers, K-Means Clustering, Boosting, Decision Tree, Random Forest (RF), Neural networks, Bayesian Networks, etc.

In our case research, 4 classifiers were finally chosen to test which could best predict users’ anger and confusion. Logistic Regression, Random Forest, Support Vector Machine and XGBoost were used twice to calculate users’ anger and confusion prediction scores.

1. Logistic Regression (LR)

Logistic Regression is the iterative display of the most powerful linear combination of variables, most likely to determine the observed outcome. (Yuvalı, Yaman & Tosun, 2022). It explores the link between independent and dependent variables. It is used for classification rather than estimation, even though the algorithm's name is regression. (Cihan et al., n.d.). Comparative Performance of Machine Learning Algorithms in Cyberbullying Detection: Using Turkish Language Preprocessing Techniques.) What is more, it aids in determining the likelihood that a new instance belongs to a specific class. Since it is a probability, the result will be somewhere between 0 and 1. As a result, in order to employ the LR as a binary classifier, a threshold must be specified to distinguish between two classes. The LR model can be used to model a categorical variable with more than two values. Studying the relationships between a set of labeled data, it helps categorize data into discrete classes. Logistic regression is one of the most frequently used methods in statistics and discrete data analysis. Jet, A., & O, H. J. (2017). Supervised Machine Learning Algorithms: Classification and Comparison. International Journal of Computer Trends and Technology, 48. http://www.ijcttjournal.org

2. Random Forest (RF)

Random Forest can handle large datasets with automatic variable selection and many estimators. It is reported to provide unbiased Jet, A., & O, H. J. (2017). A random forest is an ensemble classifier which is utilized to combine the expectations from different machine learning calculations together to create exact results. It is a combination of numerous decision trees. The default hyperparameters of RF gives great result and it is incredible at avoiding overfitting (Pretorious, Bierman & Steel, 2016) https://www.researchgate.net/publication/312486161_A_meta-analysis_of_research_in_random_forests_for_classification.

3. Support Vector Machine (SVM)

The SVM method is used to categorize linear and non-linear linear data. In short, the algorithm works as follows: it uses a non-linear matching technique for transforming the original set of data for education in a higher dimension. In this new dimension seeks the optimal linear divider super-plane, that is, the boundary that separates the blocks that belong to a class from those of a different class. With appropriate non-linear mapping to a sufficiently high dimension the data set belong to two different classes can always be separated by one superficial (Τσιλιγιάννη, 2015).

4. XGBOOST

It is a decision-tree community learning algorithm. The most contrast from other calculations lies in its versatility, which empowers quick learning through parallel and disseminated computing and gives effective memory utilization. It is free from over fitting and bias. Since it was revealed in 2016 it is considered as recent (Cihan et al., n.d.)

4.2 Modeling

As it was already mentioned, to find out whether users' anger and confusion can be predicted from mouse movements on a gamification task, we used four machine learning algorithms. Logistic Regression, Random Forest Classifier, Support Vector Machine, and XGBoost.

In the following sections, it is presented the procedure followed to find out whether and to what extent mouse movements can predict the user's emotions on a gamified task.

4.2.1 Data splitting

Once all the necessary steps have been completed for our final dataset to be built, the next step is to split the data into train and test. Data splitting is an important step in machine learning. To do so, we used the 'train_test_split' importing sklearn library. The data with the independent variables is labeled X, whereas the data with the dependent variable is labeled Y. The test size variable specifies the ratio in which the data will be split. In our case, the final dataset was divided into 70% training and 30% test data sets. As for the random_state, it regulates how the data is shuffled before the split is done. Setting the random state to a constant ensures that the same sequence of random integers is generated each time the code is executed.

```
# Create Train & Test Data
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=101)
```

Figure 26 Train and Test data

4.2.2 Accuracy Scores

- ***Logistic Regression Model***

First, LogisticRegression function was imported using sklearn.linear_model class in python. Secondly, an instant of the Model was created. Then, our model was being trained on the training data set, saving the information learned from the data.

```
# Running logistic regression model
from sklearn.linear_model import LogisticRegression
model = LogisticRegression()
result = model.fit(X_train, y_train)
```

Figure 27 Running Logistic Regression Model

With the trained model “confusion” and “anger” was predicted. The results were saved in “prediction_test”. Then, the accuracy score was measured.

```
from sklearn import metrics
prediction_test=model.predict(X_test)
# Print the prediction accuracy

print (metrics.accuracy_score(y_test, prediction_test))
```

```
0.9090909090909091
```

Figure 28 Logistic Regression Accuracy Score for Confusion

For our binary classification problem, our model predicted the right outcome in 90% of the cases. Regarding the prediction of anger, the accuracy score reached 87%.

```
from sklearn import metrics
prediction_test=model.predict(X_test)
# Print the prediction accuracy

print (metrics.accuracy_score(y_test, prediction_test))
```

```
0.8787878787878788
```

Figure 29 Logistic Regression Accuracy Score for Anger

- ***Random Forest***

The second Machine Learning prediction model was created using Random Forest Classifirer. For Algorithm, sklearn.ensemble class was used and RandomForestClassifier was imported, since we had a Classification problem. In this model, the accuracy score was 87% for “Confusion” and 81% for “anger”.

```
# random forest



from sklearn.ensemble import RandomForestClassifier
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)
model_rf = RandomForestClassifier(n_estimators=1000 , oob_score = True, n_jobs = -1,
                                  random_state =50, max_features = "auto",
                                  max_leaf_nodes = 30)
model_rf.fit(X_train, y_train)

# Make predictions
prediction_test = model_rf.predict(X_test)
print (metrics.accuracy_score(y_test, prediction_test))
```

```
0.8787878787878788
```

Figure 30 Random Forest Accuracy Score for Confusion

```
# random forest


from sklearn.ensemble import RandomForestClassifier
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)
model_rf = RandomForestClassifier(n_estimators=1000 , oob_score = True, n_jobs = -1,
                                  random_state =50, max_features = "auto",
                                  max_leaf_nodes = 30)
model_rf.fit(X_train, y_train)

# Make predictions
prediction_test = model_rf.predict(X_test)
print (metrics.accuracy_score(y_test, prediction_test))
```

```
0.8181818181818182
```

Figure 31 Random Forest Accuracy Score for Anger

- ***Support Vector Machine***

For the third Model sklearn.svm was used importing SVC. In this research for SVM model 'linear' kernel was selected (Wadikar, 2020). Both confusion accuracy score and anger accuracy score reached 87%.

```
from sklearn.svm import SVC

model.svm = SVC(kernel='linear')
model.svm.fit(X_train,y_train)
preds = model.svm.predict(X_test)
metrics.accuracy_score(y_test, preds)
```

```
0.8787878787878788
```

Figure 32 SVM Accuracy Score for Confusion

```python
from sklearn.svm import SVC

model.svm = SVC(kernel='linear')
model.svm.fit(X_train,y_train)
preds = model.svm.predict(X_test)
metrics.accuracy_score(y_test, preds)
```

```
0.8787878787878788
```

Figure 33 SVM Accuracy Score for Anger

XGBoost

The last Machine Learning Model, XGBoost. XBClassifier was impoted using xgboost. Secondly, an instant of the Model was created. Next, our model was being trained on the training data set, saving the information learned from the data. Finally, the accuracy score for the confusion prediction reached 81% while for anger 84% .

```python
from xgboost import XGBClassifier
model = XGBClassifier()
model.fit(X_train, y_train)
preds = model.predict(X_test)
metrics.accuracy_score(y_test, preds)
```

```
[19:06:26] WARNING: ../src/learner.cc:1095: Starting in XGBoost 1.3.0, the default evalua
tion metric used with the objective 'binary:logistic' was changed from 'error' to 'loglos
s'. Explicitly set eval_metric if you'd like to restore the old behavior.

0.8181818181818182
```

Figure 34 XGBoost Accuracy Score for Confusion

```
#xgboost
from xgboost import XGBClassifier
model = XGBClassifier()
model.fit(X_train, y_train)
preds = model.predict(X_test)
metrics.accuracy_score(y_test, preds)
```

```
[19:11:42] WARNING: ../src/learner.cc:1095: Startin
binary:logistic' was changed from 'error' to 'loglo
```

```
:
0.8484848484848485
```

Figure 35 XGBoost Accuracy Score for Anger

5 Results

Our findings exposed that some players completed only the first gamified task. The next figure depicts the distribution of 2 games. It is obvious that some users completed only the first gamified task and did not continue to the second, since the distribution of the first game (puzzle) is higher.

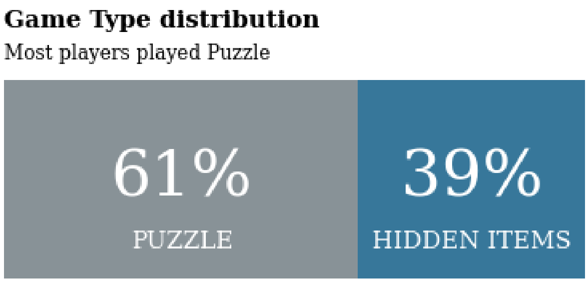

Figure 36 Game type Distribution

The next figure displays the rate of users who answered that felt confused with a percent of 31,2% and those who did not (68,8%). At the same time, 75,2% of users stated that they felt angry, while 24,8% of them did not state it.

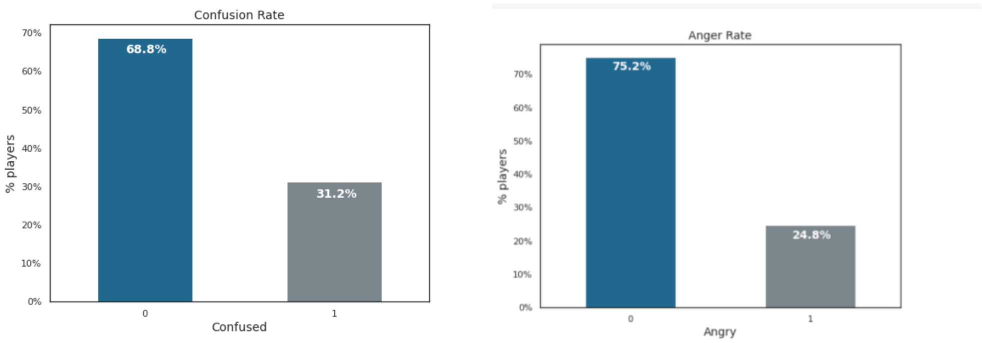

Figure 37 Confusion and Anger rate among users for the whole experiment

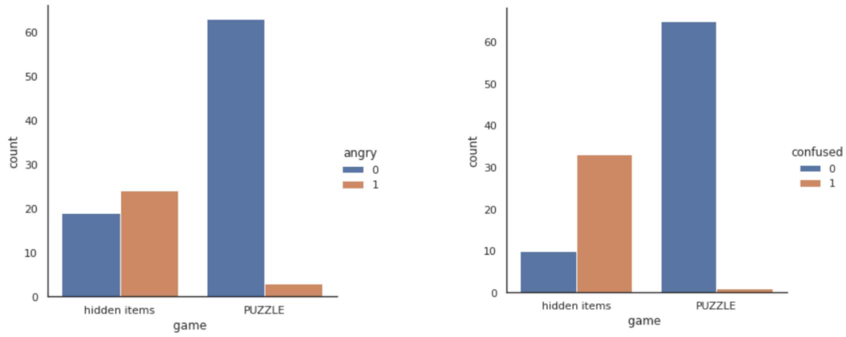

Figure 38 The number of user's who felt or not angry and confused in both games.

From the results it is clear, that in both games, more users answered that they did not feel confused or angry.

As for the mouse variables, the next two figures show the correlation between them. A -1 to +1 correlation exists. The stronger the relationship, the closer the correlation is to one; in other words, as one increases, the other increases as well, and the closer the correlation is to one, the stronger it is. If the correlation is closer to -1, one variable will fall as the other rises, rather than both rising at the same time.

In the case of “angry” target variable, there is a positive correlation with: confused_yes (if someone felt confused), game_hidden_items (our first gamified task), mean_acceleration, and mean_times_since. On the other hand, there is a negative correlation between “angry” target variable and the other variebles: var_speed, std_speed, mean_speed, confused_no (if someone stated that they did not feel confused) and our second gamified task (Puzzle).

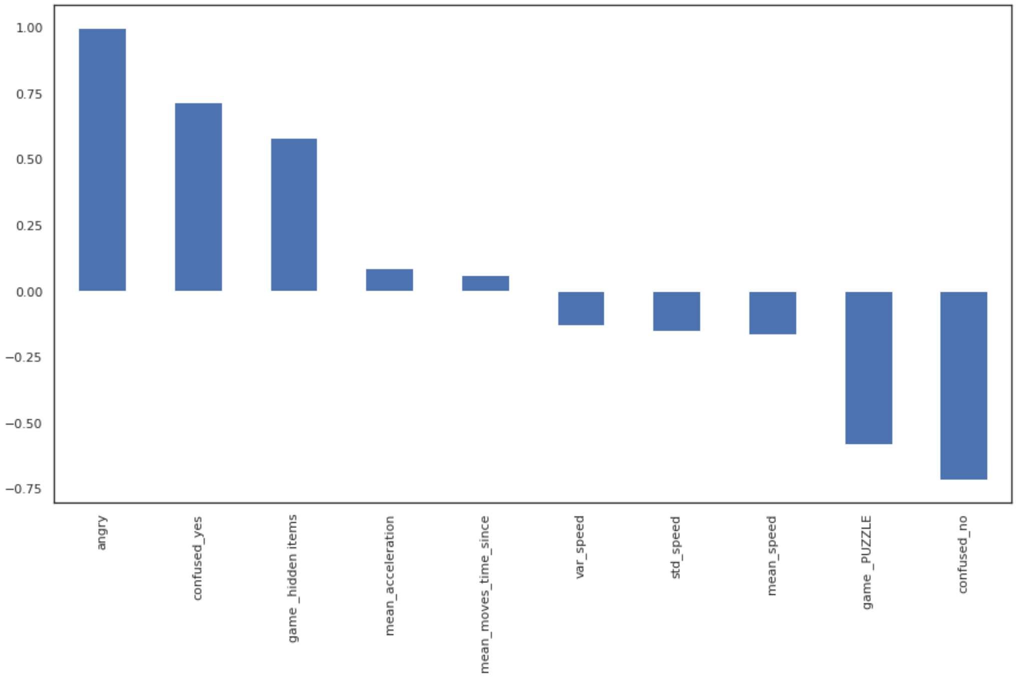

Figure 39 Correlation between variables (angry Notebook)

As for the correlation of “confused” with other variables, as it is shown in Figure 40, there is a positive correlation between “confused” target variable and: game_hidden_games, angry_yes (if someone stated that felt anger),mean_moves_times_since, mean_acceleration. However, there is a negative correlation between “confused” and the other variables.

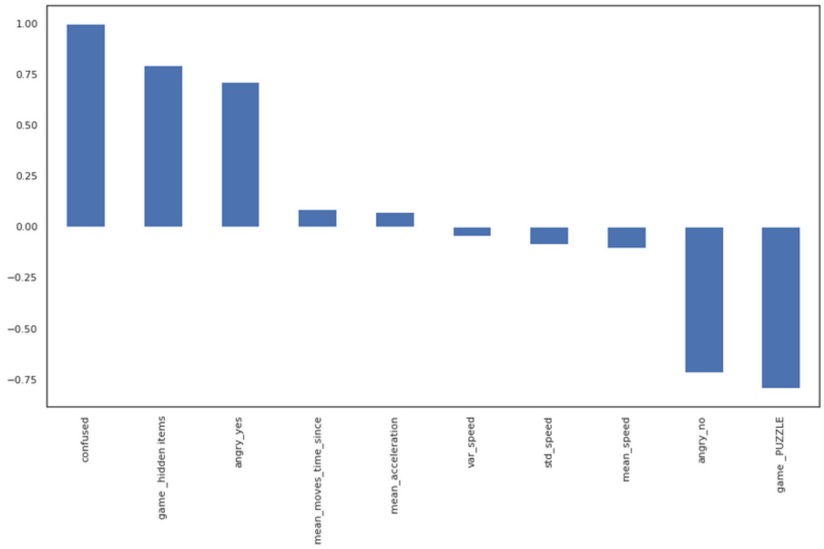

Figure 40 Correlation between variables (confusion Notebook)

The same conclusions apply from the following correlation heatmaps, that presents the correlation between the variables. Each square represents the correlation between the variables on each axis.

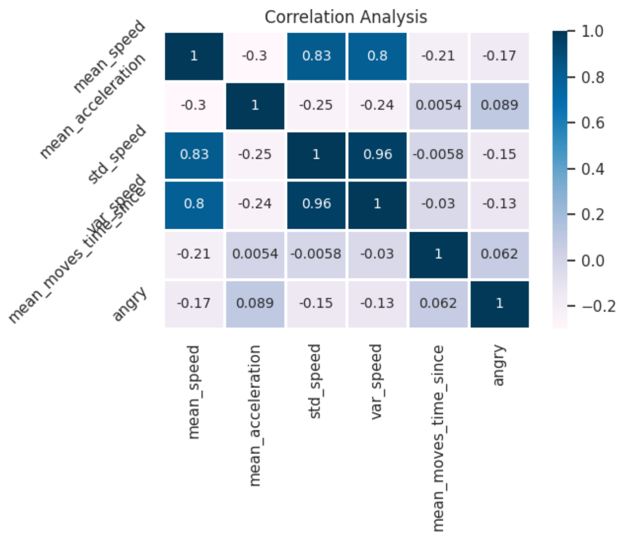

Figure 41 Correlation heatmap (angry)

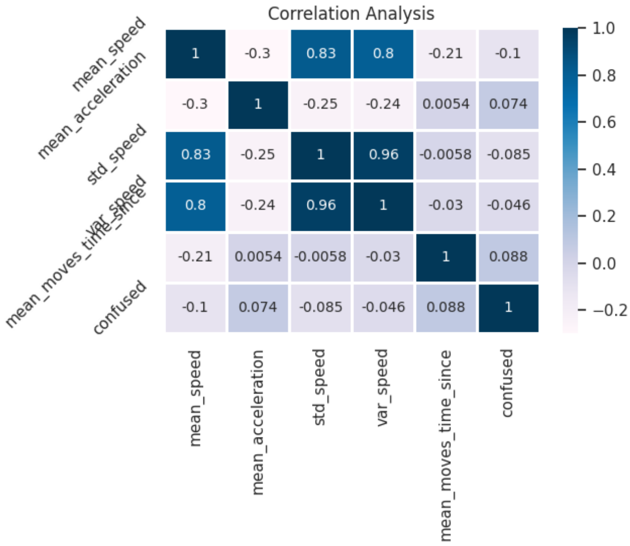

Figure 42 Correlation heatmap (confusion)

Four Machine Learning Algorithms were applied to investigate to what extend users' anger and confusion could be predicted.

Results are displayed in the next table which depicts and compares the accuracy scores for anger and confusion results concluded from Machine Learning Algorithms. Our results show that both of our Classifiers were able to predict users feeling on a gamification task. However, every model except XGBoost, achieved to predict users' confusion better than users' anger. XGBoost was the only classifier that performed better accuracy scores for users' anger than confusion. What is more, in the case of Support Vector Machine both of accuracy scores reached 87%.

Table 4 Machine Learning Algorithms and accuracy scores

Machine Learning Algorithms	Confusion Accuracy Score	Anger Accuracy Score
Logistic Regression	90%	87%
Random Forest	87%	81%
Support Vector Machine	87%	87%
XGBoost	81%	84%

6 Discussion

6.1 Principal Findings

A major aim of this study was to determine whether mouse movements can predict users' feelings when they are engaged in a gamified task, using four machine learning algorithms. The data collected from 111 users who took part in 2 gamified tasks supported that theory. The predicting scores calculated were satisfying enough to confirm that predictions can be formed from mouse movements.

Moreover, this study has also shown that a user's feelings toward a particular situation can have a significant impact on their mouse cursor moves. The hihger difficulty of the task , the higher the level of anger and confusion were

Based on the conclusions reached in this analysis, it appears that users' anger and confusion, can be predicted using mouse data gathered from a gamified environment.

Mouse's behavior varies depending on the user's emotions; for example, the pressure exerted on the mouse can increase in combination with the accumulation of the user's frustration (Yamauchi & Xiao, 2018).

In this study, the levels of anger and confusion in the case of the problematic task were greater. Thus, our study supports that theory as well.

6.2 Comparison with previous work

Previous works have clearly shown that mouse movements are directly related to user behavior and emotions (Banholzer et al., 2021; Tzafilkou & Protogeros, 2018).

Also, various algorithms have been applied to predict users' behavior or emotions through cursor movements.

However, these cases are few and in some of them, the accuracy scores were not what expected (Dias Da Silva & Postma, 2020). Since very few studies have combined mouse tracking and gamification (Betz et al., 2020), this study could be considered as one of the first that has attempted to predict users' emotions using information from cursor's movements on a gamification task. What is more, since the combination of mouse tracking with gamification has not been studied in the past, there has not been much progress in applying machine learning approaches to predict user emotions through mouse data in gamified campaigns.

6.3 Limitations

This study has also limitations and there are several things that need improvement.

- The main limitation was the lack of mouse information related to mouse clicks. Previous studies used mouse clicks in their research and drew important conclusions from them. Our work did not include mouse clicks. Consequently, the mouse features were fewer than those that we initially considered.
- Secondly, while in several cases the researchers also investigated the characteristics of the participants, in our case it did not happen. It would be interesting if we had examined information like gender, age, etc.
- Third, the lack of free time of the users decreased the number of people who managed to play the games. A bigger sample could give better results.
- Furthermore, the fact that users joined our experiment remotely, may have affected the way in which they completed the task.

Those three limitations may have an impact on our results and may had an effect our results.

6.4 Future work

Several fundamental issues in this thesis have been addressed, and these provide the direction for future research. In this section, we provide a brief overview of some research areas of future interest.

One is the potential application of gamification and mouse tracking. This combination could be useful in different fields such as education, marketing, and work environments. Predicting users ’emotions in this field can be helpful in many issues related to user behavior and mental state.

Secondly, future works can improve their sample by including more participants from different age groups with different educational backgrounds and occupations. Of course, the lack of information from clicks is an element of improvement for other researchers. Furthermore, it would be great for interesting positive emotions to be predicted as well.

One great example for future application could be the prediction of students' or workers' stress and anger during their exams or education. After COVID, the procedures followed in training and work have changed radically. It has been observed that new procedures have caused great stress to students and employees. Gamification could be an advantageous method in online learning at schools and companies. This idea, in combination with the detection of mouse movements and the prediction of users' emotions, could help in the reduction of negative emotions and the improvement of the way that these processes work.

6.5 Implications

The main aim of this study was to invest if it is possible to predict a user’s mental state by tracking the activity of their cursor.

We have done so by applying mouse tracking on users who were engaged in a gamified task, and we collected data which were finally fed to 4 machine learning algorithms.

Accordingly, the first major practical contribution of the present research is to provide data and information both to gamification and mouse tracking fields, with special attention to their combination.

This information is important given that only few studies have analyzed a similar topic in the past. What is more, this research presented and analyzed the concepts of gamification and how its benefits in other fields, mainly in marketing.

Besides demonstrating the benefits of gamification, this study also suggested specific implications where gamification and mouse tracking could be beneficial. Some of these implications are:

- First, in the field of marketing, the combination of a gamified strategy and mouse tracking could add great value. Considering how different levels of emotions lead to different purchasing behaviors, the prediction of emotions could be used by marketers to improve their campaigns and increase their sales.

- Second, in education, the prediction of specific emotions among students could also affect their performance. For example, tracking the activity of students’ cursors in the lesson process or a test which is could help educators to improve their teaching method or to identify problematic student behaviors like copying and prevent them. Furthermore, applying techniques of gamification on tests and teaching it could lead to better results.

- Third, in online gaming, mouse tracking could improve players’ experience by helping the creators to understand and identify which parts of the games are more fun and attract players' attention more.

- Finally, In the field of web development, mouse tracking in conjunction with game elements can improve the user experience as well by helping designers and developers to detect the points which are more preferable for users and which are not.

7 Conclusion

Taken together, the findings of the current study, it is revealed that the mouse cursor can be used to predict users' feelings, such as anger and confusion, while a user is engaged in a gamification task. Thus, it demonstrates that the different levels of feeling that was arisen when someone interacts with a gamified task can be predicted using machine learning algorithms.

Specifically, two gamified tasks were created using the Easypromos application. One of our two tasks was problematic. Thus, different levels of anger and confusion were evoked. A mouse tracking mechanism recorded users' activity. This information was processed in Jupiter Notebook using Python and its' Library. Finally, nine mouse features, including: player, mean_speed, mean_acceleration, std_speed, var_speed, mean_moves_time_since, ungry, confused, and the type of each game, were created for each participant and stored in a CSV file which was our final data set.

The next step was to upload and read the CSV file using the Kaggle application. Again, using Python, the data set was processed and reached the appropriate format to feed four Machine Learning models, Logistic Regression, Random Forest, Support Vector Machine, and XGboost. Based on the results, "Confusion" can be predicted better than "Anger" both in Logistic Regression, Random Forest, and Support Vector Machine. On the other hand, XGboost's accuracy score was higher for the "Anger" variable.

The main contribution of this work is to further introduce the concept of gamification in Mouse tracking by revealing that it is possible to make important predictions about user behavior and emotions, which can be very beneficial in many other fields

References

Adamou, B. (2018). Games and Gamification in Market Research: Increasing Consumer Engagement in Research for Business Success, Kogan Page (1-24

Arapakis, I., Lalmas, M., & Valkanas, G. (2014). *Understanding Within-Content Engagement through Pattern Analysis of Mouse Gestures*. 1439–1448. https://doi.org/10.1145/2661829.2661909

Arapakis, I., & Leiva, L. A. (2016). Predicting user engagement with direct displays using mouse cursor information. *SIGIR 2016 - Proceedings of the 39th International ACM SIGIR Conference on Research and Development in* Information *Retrieval*, 599–608. https://doi.org/10.1145/2911451.2911505

Arnedo-Moreno, J., Mora, A., Riera, D., & González, C. (2017). *A literature review of gamification design frameworks*. https://www.researchgate.net/publication/279059685

Aviz, I. L., Souza, K. E., Ribeiro, E., de Mello, H., & Seruffo, M. C. da R. (2019). Comparative study of user experience evaluation techniques based on mouse and gaze tracking. *Proceedings of the 25th Brazillian Symposium on Multimedia and the Web, WebMedia 2019*, 53–56. https://doi.org/10.1145/3323503.3360623

Banholzer, N., Feuerriegel, S., Fleisch, E., Bauer, G. F., & Kowatsch, T. (2021). Computer mouse movements as an indicator of work stress: Longitudinal observational field study. *Journal of Medical Internet Research*, *23*(4). https://doi.org/10.2196/27121

Betz, Simon & Szekely, Eva & Zarrieß, Sina & Schröer, Marin & Schade, Leonie & Wagner, Petra. (2020). HESITATION PROCESSING ANALYSIS USING CONTINUOUS MOUSE-TRACKING AND GAMIFICATION. https://www.researchgate.net/publication/339457889

Chen, Y., Liu, Y., Zhang, M., & Ma, S. (2017). User Satisfaction Prediction with Mouse Movement Information in Heterogeneous Search Environment. *IEEE Transactions on Knowledge and Data Engineering*, *29*(11), 2470–2483. https://doi.org/10.1109/TKDE.2017.2739151

Christians, G. (2018). *Scholar Commons Scholar Commons The Origins and Future of Gamification The Origins and Future of Gamification*. https://scholarcommons.sc.edu/senior_theses

Cihan Ates, E., Bostanci, E., & Güzel, S. (2021). Comparative Performance of Machine Learning Algorithms in Cyberbullying Detection: Using Turkish Language Preprocessing Techniques. https://doi.org/10.48550/arXiv.2101.12718

Dias da Silva, M. R., & Postma, M. (2020). Wandering minds, wandering mice: Computer mouse tracking as a method to detect mind wandering. Computers in Human Behavior, 112, 106453. https://doi.org/10.1016/J.CHB.2020.106453

Diaz, F., White, R.W., Buscher, G., & Liebling, D.J. (2013). Robust models of mouse movement on dynamic web search results pages. Proceedings of the 22nd ACM international conference on Information & Knowledge Management. 1451-1460. 10.1145/2505515.2505717.

Giannakopoulou, S. (2021). Digital advertising and gamification technologies in affecting consumer behavior.

Grif, H.-S., & Farcas, C. C. (2016). Mouse Cursor Control System Based on Hand Gesture. Procedia Technology, 22, 657–661. https://doi.org/10.1016/j.protcy.2016.01.137

Høgenhaug, P. S. (2012). Gamification And UX: Where Users Win Or Lose . Smashing Magazine.

Huang, W.H.-Y. and Soman, D. (2013) Gamification of Education. Research Report Series: Behavioural Economics in Action. Rotman School of Management, University of Toronto, available at: https://inside.rotman.utoronto.ca/behaviouraleconomicsinaction/files/2013/09/GuideGamificationEducationDec2013.pdf

Jaiswal, A. K., Tiwari, P., & Hossain, M. S. (2020). Predicting users' behavior using mouse movement information: an information foraging theory perspective. *Neural Computing and Applications*. https://doi.org/10.1007/s00521-020-05306-7

Jet, A., & O, H. J. (2017). Supervised Machine Learning Algorithms: Classification and Comparison. *International Journal of Computer Trends and Technology*, *48*. http://www.ijcttjournal.org

Johnson, D. (2022).Python Pandas Tutorial: DataFrame, Date Range, Use of Panda, Guru99, available at: https://www.guru99.com/python-pandas-tutorial.html

Kaarlehto, K. (2020). Gamification in new customer acquisition of a digital marketing service provider. Laurea University of Applied Sciences

Kalogiannakis, M., Papadakis, S., & Zourmpakis, A. I. (2021). Gamification in science education. A systematic review of the literature. Education Sciences, 11(1), 1–36. https://doi.org/10.3390/educsci11010022

Kim, S., Song, K., Lockee, B., & Burton, J. (2018). What is Gamification in Learning and Education? In Gamification in Learning and Education (pp. 25–38). Springer International Publishing. https://doi.org/10.1007/978-3-319-47283-6_4

Kiryakova, Gabriela & Angelova, Nadezhda & Yordanova, Lina. (2014). Gamification in Education. Conference: 9th International Balkan Education and Science Conference At: Edirne, Turkey

Konstan, J. A., Chi, E. H., Höök, K. (Writer on technology design), & SIGCHI (Group : U.S.). 2012. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA.

Leiva, L. A., Arapakis, I., & Iordanou, C. (2021). My Mouse, My Rules: Privacy Issues of Behavioral User Profiling via Mouse Tracking. *CHIIR 2021 - Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*, 51–61. https://doi.org/10.1145/3406522.3446011

Lucassen, G., & Jansen, S. (2014). Gamification in Consumer Marketing - Future or Fallacy? *Procedia - Social and Behavioral Sciences*, *148*, 194–202. https://doi.org/10.1016/j.sbspro.2014.07.034

Pretorius, A., Bierman, S., & Steel, S. J. (2017, January 10). A meta-Analysis of research in random forests for classification. *2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference, PRASA-RobMech 2016*. https://doi.org/10.1109/RoboMech.2016.7813171

Rodrigues, M., Gonçalves, S., Carneiro, D., Novais, P., Fdez-Riverola, F. (2013). Keystrokes and Clicks: Measuring Stress on E-learning Students. In: Casillas, J., Martínez-López, F., Vicari, R., De la Prieta, F. (eds) Management Intelligent Systems. Advances in Intelligent Systems and Computing, vol 220. Springer, Heidelberg. https://doi.org/10.1007/978-3-319-00569-0_15

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, *69*, 371–380. https://doi.org/10.1016/j.chb.2016.12.033

Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, *2*(3). https://doi.org/10.1007/s42979-021-00592-x

Schoemann, M., O'hora, D., Dale, R., & Scherbaum, S. (2021). *Using mouse cursor tracking to investigate online cognition: Preserving methodological ingenuity while moving toward reproducible science*. https://doi.org/10.3758/s13423-020-01851-3/Published

Schöner, G., & Lins, J. (2017). *Mouse Tracking Shows Attraction to Alternative Targets While Grounding Spatial Relations*. https://www.researchgate.net/publication/316924038

Souza, K. E. S., Seruffo, M. C. R., de Mello, H. D., Souza, D. D. S., & Vellasco, M. M. B. R. (2019). User Experience Evaluation Using Mouse Tracking and Artificial Intelligence. *IEEE Access*, *7*, 96506–96515. https://doi.org/10.1109/ACCESS.2019.2927860

Stillman, P. E., Shen, X., & Ferguson, M. J. (2018). How Mouse-tracking Can Advance Social Cognitive Theory. In *Trends in Cognitive Sciences* (Vol. 22, Issue 6, pp. 531–543). Elsevier Ltd. https://doi.org/10.1016/j.tics.2018.03.012

Tanjim-Al-akib, M., Hosne-Al-walid, Ashik, L. K., & Chowdhury, K. (2017). User-modeling and recommendation based on mouse-tracking for E-commerce websites. *19th International Conference on Computer and Information Technology, ICCIT 2016*, 517–523. https://doi.org/10.1109/ICCITECHN.2016.7860252

Triantoro, T., Gopal, R., Benbunan-Fich, R., & Lang, G. (2019). Would you like to play? A comparison of a gamified survey with a traditional online survey method. *International Journal of Information Management*, *49*, 242–252. https://doi.org/10.1016/j.ijinfomgt.2019.06.001

Tzafilkou, K., & Protogeros, N. (2020). Monitoring Mouse Behavior in e-learning Activities to Diagnose Students' Acceptance Items of Perceived Usefulness and Ease of Use. *The European Educational Researcher*, *3*(1), 21–27. https://doi.org/10.31757/euer.312

Tzafilkou, K., & Protogeros, N (2018). Mouse behavioral patterns and keystroke dynamics in End-User Development: What can they tell us about users' behavioral attributes? Computers in Human Behavior, 83, 288–305. https://doi.org/10.1016/J.CHB.2018.02.012

Tzafilkou, K., Protogeros, N., & Yakinthos, C. (2014). Mouse Tracking for Web Marketing: Enhancing User Experience in Web Application Software by Measuring Self-Efficacy and Hesitation Levels. *International Journal of Strategic Innovative Marketing*. https://doi.org/10.15556/ijsim.01.04.005

Vicario, C. M., & Newman, A. (2013). Emotions affect the recognition of hand gestures. Frontiers in Human Neuroscience, 7, Article 906. https://doi.org/10.3389/fnhum.2013.00906

Xouridas Nikolaos. (2020). Gamification for Marketing Research-Measure UX Features and User Engagement on an A/B Testing Marketing Campaign.

Yamauchi, T. (2013). Mouse trajectories and state anxiety: Feature selection with random forest. Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013, 399–404. https://doi.org/10.1109/ACII.2013.72

Yamauchi, T., & Xiao, K. (2018). Reading Emotion from Mouse Cursor Motions: Affective Computing Approach. *Cognitive Science*, *42*(3), 771–819. https://doi.org/10.1111/cogs.12557

Yi, Q., Xiong, S., Wang, B., & Yi, S. (2020). Identification of trusted interactive behavior based on mouse behavior considering web User's emotions. *International Journal of Industrial Ergonomics*, *76*. https://doi.org/10.1016/j.ergon.2019.102903
Yuvalı, M., Yaman, B., & Tosun, Ö. (2022). Classification Comparison of Machine Learning Algorithms Using Two Independent CAD Datasets. *Mathematics*, *10*(3). https://doi.org/10.3390/math10030311

Τεχνολογιας Πληροφορικης Και Υπολογιστων, T., & Κ Τσιλιγιάννη, E. (n.d.). *ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ Αλγόριθμοι Classification σε Big Data ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ*.