Use of Telemetry Data and Emotion Analysis to complement the analysis of Problematic Gaming Behaviors

Beltrán Vázquez Liniers

Mestrado em Ciência de Dados

Dissertação orientada por:
Profª. Doutora Ana Paula Pereira Afonso
Prof. Doutor Manuel Joao Caneira Monteiro da Fonseca
Acknowledgements

I would like to express my gratitude to my academic advisors and mentors: Dra. Ana Paula Afonso and Dr. Manuel J. Fonseca. Their guidance, expertise, and constructive feedback have been essential in introducing me to the research area, as both are accomplished researchers. I would like to extend my heartfelt gratitude to Psychologist Joana Cardoso. Her contribution to the interpretation of the developed metrics has been crucial in the completion of this project. Also thanks to her, the entire group has learned to think from a psychological point of view, which we take with us to real life as well.

Personally, I felt inspired and motivated by my family, which has strong scientific roots, especially my father. He is an international eminence in physics and nanotechnology. But also to my mother and brother because of their daily preoccupation with myself.

Special thanks to my friends, with whom I have played Counter-Strike since I was a teenager in cyber cafes when no one had internet at home, until today when everyone plays at his one place. Particularly to Peble Pozzo and Alberto Bire Gele, my gaming counselors.
To everyone who enjoy video games with responsibility
Resumo

A perturbação de jogos da internet é uma perturbação definida pela Associação Americana de Psiquiatria (APA), em 2013, e pela Organização Mundial de Saúde (OMS), em 2018, nos seus manuais de diagnóstico (DSM-5 e ICD-11 respetivamente). As duas instituições têm os seus próprios critérios de diagnóstico para a identificação da adição aos videogames. A APA denomina a adição aos videogames com o termo Internet Gaming Disorder (IGD). Enquanto que a OMS a denomina como Gaming Disorder. A nossa investigação está orientada a seguir a definição do IGD. Os critérios para ser diagnosticado com IGD são: (1) Preocupação com os videogames, (2) sintomas de abstinência quando o videogame é retirado, (3) tolerância ou necessidade de jogar cada vez mais tempo, (4) incapacidade de reduzir o tempo de videogames ou fracassar nas tentativas, (5) abandono ou perda de interesse por outras atividades a causa dos videogames, (6) Continuar a jogar apesar de saber as consequências, (7) mentir a terapeutas ou entes queridos sobre o tempo dedicado aos videogames, (8) usar os videogames como forma para aliviar o mal-estar, e (9) ter prejudicado ou perdido uma relação pessoal ou trabalho por causa dos videogames. Para ser diagnosticado com IGD, é necessário estarem presentes 5 dos 9 critérios durante pelo menos 12 meses.

Apesar destas definições, os profissionais de saúde mental continuam a encontrar dificuldades no processo de identificação desta perturbação mental. Por outro lado, há uma falta de consenso entre os profissionais de todo o mundo devido a existir uma grande variabilidade na interpretação destes critérios e no tratamento que deve ser efetuado para cada paciente.

As atuais ferramentas de deteção de adição aos videogames baseiam-se, na sua maioria, em questionários sobre a autoavaliação. Assim, os pacientes preenchem questionários compostos pelos critérios de diagnóstico do DSM-5, em que respondem sobre o seu tempo de jogo, como se sentem quando não estão a jogar, e a relação entre a sua atividade de jogo e os seus familiares próximos e terapeutas. A partir daqui, surge outra preocupação das metodologias atuais: a informação obtida é subjetiva. Não existe um controlo de horas, em consequência os jogadores não se apercebem do tempo dispendido no jogo, além de não recordarem exatamente como se sentiram.

Existem múltiplas metodologias para o tratamento da adição aos videogames, algumas proveem medicamentos farmacológicos, e outras baseiam-se em Terapias Cognitivo-Comportamentais (TCC), ou uma combinação de ambas. Na avaliação e durante o processo terapêutico, os clientes podem preencher questionários de autorrelato de forma a monitorizar os sintomas da perturbação. Estes questionários tendem a abranger todos os sintomas da IGD e comportamentos ou problemáticas
relacionados. Deste modo o terapeuta pode analisar rapidamente como e quando jogou o paciente nos últimos dias.

O objetivo desta proposta é ajudar a clarificar os critérios de diagnóstico e monitorização da terapia. Concretamente, o fim é definir métricas e visualizações relevantes que possam funcionar como ferramentas de diagnóstico e apoiar a monitorização da terapia. Estes indicadores são construídos utilizando a telemetria do jogo, e dados sobre as emoções recolhidas a partir de gravações de vídeo do rosto do jogador durante o jogo. Os KPIs são definidos para obter os dados mais objetivos possíveis, pelo que também podem ser úteis para analisar se o paciente tem uma boa compreensão do seu verdadeiro comportamento de jogo.

Os métodos para desenvolver a solução passam pela procura de uma fonte de dados aberta e de qualidade. Após múltiplas pesquisas a API de FACEIT é a que fornece os melhores dados de telemetria em termos de usabilidade e acessibilidade. Além disso, utilizamos o sistema de reconhecimento de emoções DeepFace para extrair os sentimentos do jogador a partir da gravação, também de maneira open-source. As principais características recolhidas são o tempo passado a jogar e os níveis de cada emoção primária (raiva, felicidade, nojo...) por segundo. Uma vez recolhidos os dados, desenvolvemos múltiplas visualizações que podem ajudar os terapeutas em ambas as perspectivas, tendo uma visão histórica do comportamento de jogo, para uma fase inicial do tratamento; e um comportamento de jogo muito detalhado dos últimos dias, para uma análise rápida, atualizado antes de cada consulta.

Para a visualização na fase inicial do tratamento desenvolvemos uma série de gráficos que mostram o comportamento do jogador durante dois anos, 2021 e 2022, a partir dos dados telemétricos. Estes diagramas fazem diferenciação entre os tempos de jogo durante o fim-de-semana e no meio da semana. Também segregam por horas dedicadas de manhã, tarde e noite. Estas informações podem ser muito relevantes quando o terapeuta tenta analisar se o gaming está a substituir outras atividades. Uma característica que é calculada ao início é a sessão. Uma sessão é composta por uma ou várias partidas que distam não mais do que 30 minutos entre elas. De maneira que podemos saber se joga todas as partidas consecutivamente ou se faz pausas. É também muito interessante conhecer quantas sessões joga por mês ou semana. Segundo a duração e a frequência das sessões podemos concluir se um gamer joga de maneira mais ou menos intensa que outro. Sessões mais compridas podem ajudar a determinar se o jogador tem dificuldades para parar de jogar.

Nas métricas desenhadas para a segunda perspectiva, visualização prévia a consulta, analisamos também o tempo de jogo e a intensidade, conjuntamente com as emoções sentidas nesses últimos dias. Assim, são medidas o número de horas jogadas, o número e duração das sessões, e os níveis de tristeza, felicidade ou medo. Permitindo também analisar a evolução da última semana e aos últimos 30 dias, o qual dá uma visão ao terapeuta que lhe permite saber sobre a evolução e o sucesso do processo terapêutico.

Uma das preocupações dos psicólogos é saber como os videogames interferem no estado anímico do jogador. De maneira que possamos saber se o jogador está zangado por causa do jogo ou se já sentia mal-estar antes de começar, utilizando o gaming como via de escape, o que
poderia ser um sintoma de IGD. Para isso, desenvolvemos um gráfico que mostra as emoções sentidas nos primeiros e últimos cinco minutos de cada match, deste modo comparamos os níveis de medo e alegria, por exemplo, ao início e ao final, e analisamos se o mood do jogador mudou durante o *match*.

A interpretação dos indicadores e visualizações desenhadas permitem avaliar a presença de alguns dos critérios de diagnóstico, já que se o jogador dedica muitas horas aos videogames, além das suas atividades profissionais, é possível que não possa dedicar o tempo suficiente aos relacionamentos pessoais, o que pode responder ao critério número 9 do DSM-5. Por outro lado, se houver muita diferença entre o respondido pelo paciente e o analisado nas métricas podemos concluir que há discrepâncias entre o relatado nos questionários e o analisado nas métricas, que poderá significar a presença do critério 7.

Para a análise dos resultados, baseámos as nossas conclusões em 4 participantes com características diferentes selecionados a partir de uma comunidade de jogadores portuguesa chamada Hub SAW Portugal. Neste documento, exploramos todas as métricas desenvolvidas, bem como os gráficos e diagramas que analisam os diferentes comportamentos de cada participante. Os resultados são encorajadores, as métricas são bastante diferentes entre os quatro gamers, pelo que poderão ser úteis no diagnóstico da IGD e na monitorização do tratamento.

Contudo, a solução proposta pode contribuir muito favoravelmente na identificação, diagnóstico e tratamento de pessoas com perturbações devidas à adição aos videogames. É uma solução que responde com garantias a várias das preocupações dos terapeutas dedicados ao IGD.

**Palavras-chave:** Perturbação de jogos da Internet, Dados telemetria, Reconhecimento emoções por imagens, Ciência de dados, Psicologia
Abstract

Gaming addiction is a perturbation mentioned by the American Psychiatric Association (APA), in 2013, and World Health Organization (WHO), in 2018, in their publications (DSM-5 and ICD-11, respectively). The two institutions have their own diagnostic criteria for the identification of the gaming addiction. The APA uses the term Internet Gaming Disorder (IGD) meanwhile WHO uses Gaming Disorder (GD). This investigation is oriented to the IGD definition. Some of the identification and diagnosis criteria for IGD are: Withdrawal symptoms when gaming is taken away or not possible, tolerance, deceiving family members or others about the amount of time spent on gaming, or having jeopardized a job or relationship due to gaming. Despite these definitions, mental health professionals still encounter difficulties in the process of identification of this mental disorder. On the other hand, there exists a lack of consensus among professionals all around the world because there is high variability in the interpretation of these criteria and the treatment that should be carried out for each patient.

The current screening tools are mostly based on self-related reports. Hence, the patients fulfil a questionnaire answering about their playtime, how they feel when they are not playing, and the relationship between their gaming activity and their close relatives and therapists. From here gives off another concern of the current methodologies, the gathered information is subjective. In addition, it is difficult for the patients to remember exactly how long time they spent playing.

The objective of this proposal is to clarify and unify diagnosis and therapy monitoring criteria. Particularly, the goal is to define relevant metrics and visualizations that could assist with the current diagnostic tools, and support the therapy monitoring. These indicators are built using gaming telemetry, and data on emotion gathered from video recordings during the games. The KPIs are defined to get the most objective data possible, hence they also can be helpful to analyze whether the patient has a good understanding of his real gaming behaviour.

The methods to develop the solution pass through the finding of an accurate and open data source, after a thorough search the best option is the API of FACEIT. Furthermore, we use the DeepFace image emotion recognition system to extract the feelings of the player. Also open-source. The main features collected are time spent playing and the levels of every primary emotion (Anger, happiness, disgust...) per second. Once data is collected, we developed multiple visualizations that can help therapists in both perspectives, having a historical overview of gaming behavior, for an initial phase of the treatment; and having a very detailed gaming behavior of the last days, for a quick analysis, updated before each consultation.

For the analysis of the results, we based our conclusions on 4 participants with different characteristics. In this document, we explore all the metrics developed as well as the charts and plots
analyzing the different behaviors of every participant. The results are encouraging, since the metrics show differences between the participants. Hence they could be useful in the diagnosis of IGD and the treatment monitoring.

In conclusion, the proposed solution can contribute to the identification, diagnosis and treatment of people with possible addictive gaming behaviors. It is a solution that can help overpass therapists’ concerns when recognizing IGD.

**Keywords:** Gaming disorder, Gaming telemetry, Video emotion recognition, Data Science, Psychology
Glossary

**ADHD**  Attention Deficit Hyperactivity Disorder. 11, 12

**APA**  American Psychiatric Association. v, ix, 2, 8, 12

**API**  Application Programming Interface. vi, ix, 5, 19, 36, 38–40, 44, 65

**ASD**  Autism spectrum disorder. 11, 12

**AUD**  Alcohol Use Disorder. 14

**BAS**  Behavioral Activation System Scales. 26

**BIS**  Behavioral Inhibition System. 26

**BIS-11**  Barratt Impulsiveness Scale version 11. 14

**CBT**  Cognitive-Behavioural Treatment. 10, 32

**CNN**  Convolutional Neural Network. 27, 41

**CS**  Counter Strike. 18, 19, 38

**CSE**  Core Self Evaluation. 10

**DL**  Deep Learning. 27

**DSM-5**  Diagnostic and Statistical Manual of Disorders. v, ix, 2, 8, 10, 11, 17, 22, 23, 26, 34, 59

**ECG**  electrocardiogram. 23

**EEG**  Electroencephalography. 23, 24, 26

**EMG**  electromyography. 23

**EREC**  Emotion RECognition. 25

**ERP**  Event Related Potential. 24

**FER**  Face Emotion Recognition. 40, 41
fNIRS  functional near-infrared spectroscopy. 27

FPS  First Person Shooter. 17–19

GASA  Gaming Addiction Scale for Adolescents. 22

GD  Gaming Disorder. ix, 2

GSR  galvanic skin response. 23, 24

GUR  Game User Research. 23

HLTV  Half-life Television. 38

HRV  Heart Rate Variability. 23, 24

ICD-11  International Classification of Diseases. v, ix, 2, 8, 47

IFG  Inferior Frontal Gyrus. 27

IGD  Internet Gaming Disorder. v, vii, ix, x, xix, 2–5, 7–12, 14, 16–18, 20–29, 31–35, 47, 57, 59–61, 63–65

IGDS  Internet Gaming Disorder Scale. 21

IGDS9-SF  Internet Gaming Disorder Scale 9 questions - Short Form. 5, 21, 22, 34, 35, 47, 57, 60

IMCO  Internal Market Consumer Protection. 15

KPI  Key Performance Indicator. vi, ix, 4, 35, 36, 44, 53, 57, 59, 61, 63

LLP  Low Late Potential. 24

MK-SVM  Multi Kernel-Support Vector Machine. 26

ML  Machine Learning. 26, 27, 29, 63, 64

MMORPG  Multi Massive Online Role Playing Games. 17, 18

MOBA  Multiplayer online battle arena. 17, 18

OCD  Obsessive Compulsive Disorder. 11

PET  Positron Emission Tomography. 26
**PIPATIC** Programa Individualizado Psicoterapéutico para la Adicción a las Tecnologías de la Información y la Comunicación (Individualized Psychoterapeutic Program for Information and Communication Technologies Addiction). 10, 31, 32

**POGQ** Problematic Online Gaming Questionnaire. 21

**PPG** photoplethysmography. 23, 24

**RF** Retina Face. 41

**TCC** Tratamento Cognitivo Comportamental. v

**VAT** Video Game Addiction Test. 22

**WHO** World Health Organization. ix, 2, 8, 47
List of Figures

2.1 Visualization of the emotion order in function of the polarity and intensity [1].

3.1 Hours spent on playing video games in hours per week. Source Statista.

4.1 Gaming data life cycle. The figure is reproduced from Seif El-Nasr et al. (2013) [2].

4.2 Face detector testing - Average running time in seconds.

4.3 Emotion raw data example. Each row in the table represents the percentage time felt per emotion and second.

4.4 Percentage time felt, by Participant 4, during one round per emotion and second. Highlighted three events, two kills and one round end.

4.5 Percentual emotion average felt during each event.

4.6 Percentual emotion average felt during 5 matches of Participant 3. Separated by sessions.

4.7 Percentual emotion average felt on the first and last 5 minutes during 5 matches of Participant 3.

4.8 Main yearly telemetry metrics for the 4 participants.

4.9 Participant 2 total monthly hours during afternoon and evening.

4.10 Participant 3 total monthly hours during afternoon and evening.

4.11 Participant 3 total monthly hours during midweek and weekend.

4.12 Participant 3 total monthly hours by year.

4.13 Participant 2 monthly number of matches vs number of sessions.

4.14 Participant 4 monthly number of matches vs number of sessions.

4.15 KPIs related to the gaming time during the last 7 days, compared with the previous week and month.

4.16 KPIs related to the intensity and duration of the sessions during the last 7 days, compared with the previous week and month.

4.17 Participant 1 daily number of matches vs number of sessions, during the last 2 weeks.

4.18 Participant 4 daily number of matches vs number of sessions, during the last 2 weeks.

4.19 Percentual average emotion felt by Participant 4 during the last 2 weeks.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.20</td>
<td>Percentual emotion average felt during 5 matches of Participant 3. Separated by sessions</td>
<td>56</td>
</tr>
<tr>
<td>4.21</td>
<td>Percentual emotion average felt on the first and last 5 minutes of each match during the last session (4 matches). Participant 4</td>
<td>56</td>
</tr>
<tr>
<td>A.1</td>
<td>raw data columns extracted from FACEIT API</td>
<td>79</td>
</tr>
<tr>
<td>A.2</td>
<td>raw data columns extracted from FACEIT Demofiles</td>
<td>80</td>
</tr>
<tr>
<td>A.3</td>
<td>raw data columns extracted from video files through DeepFace Python library</td>
<td>81</td>
</tr>
<tr>
<td>A.4</td>
<td>No smoothed emotions, Participant 4</td>
<td>81</td>
</tr>
</tbody>
</table>
Contents

Glossary xvi

List of figures xviii

1 Introduction 1

1.1 Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

1.2 Goals . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4

1.3 Contributions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5

1.4 Structure of the document . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5

2 Background 7

2.1 Internet Gaming Disorder . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7

2.2 Risk factors in IGD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8

2.3 Comorbidities in IGD . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

2.4 Emotions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12

2.5 IGD compared with gambling disorder . . . . . . . . . . . . . . . . . . . . . . . 14

2.6 Gaming motives . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15

2.7 Gaming genres . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

2.8 Counter-Strike, Steam, FACEIT and Twitch . . . . . . . . . . . . . . . . . . . . 18

2.9 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20

3 Related work 21

3.1 Screening tool questionnaires . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21

3.2 Approaches based on physiological measures . . . . . . . . . . . . . . . . . . . . 23

3.3 Approaches based on emotion recognition . . . . . . . . . . . . . . . . . . . . . 24

3.4 Approaches based on neuroimaging data with Machine Learning . . . . . . . . 26

3.5 Game telemetry analysis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27

3.6 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29

4 Process of knowledge discovery through game data 31

4.1 Current intervention protocol . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31

4.2 Definition of the proposed solution . . . . . . . . . . . . . . . . . . . . . . . . . 34
Chapter 1

Introduction

In this chapter, we present the project’s motivation, outlining the driving factors behind its inception. Furthermore, we delineate the main objectives that serve as guiding principles throughout the document, along with the anticipated outcomes aimed at achieving substantial progress and impact.

1.1 Motivation

Over the last few decades, video games are a major human activity that has become increasingly more complex, diverse, realistic, and socially interactive, and are played by billions of people across geographical areas, cultures and demographics [2]. In Europe, a recent report [3] concludes that games are played by more than two-thirds of children and adolescents, and a substantial number of adults also play games. Amateur, professional, and spectator players alike are all interested in e-sports (electronic sports) and video games.

Several studies have been undertaken to analyze the benefits and negative consequences of playing video games on an individual’s physical, mental, and social well-being [4]. A vast majority of these studies have been focused on the negative effects of video games, related to addiction, depression, aggression, and physical health problems, among others [5, 6, 7, 8]. Nevertheless, clinical implications could not always be negative. Positive outcomes can arrive if avoiding off-line life makes the gamer feel better and implications are used briefly. Analogously, if this mechanism becomes persistent and compromises other personal life activities, they can be considered problematic and pathological gamers [9]. Griffiths analyzed two gamers with the same gaming time declared. Despite behaving similarly during gameplay, they differed in their psychological motivation and how they perceived and experienced gaming in their lives. He concludes that the same gaming behavior does not influence in the same way [10]. However, these research lines have received much less attention than those which analyze negative consequences. Both of the gamers in this study claimed to be playing for up to 14 h a day yet and although they were behaviorally identical in terms of their game playing, they were very different in terms of psychological motivation and the meaning and experience of gaming within their lives.

Many other studies pointed out the benefits of playing [11], and suggest that video games en-
hance cognitive abilities [12], improve socialization, provide stress relief, can have an educational value [13], can help players improve problem-solving skills, and can improve hand-eye coordination and fine motor skills [14].

Internet Gaming Disorder (IGD) appeared in the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders) in 2013, in the section of conditions for future study, with the intention of motivating discussion and research on it [15]. This manuscript is assigned by the American Psychiatry Association (APA). With its origin and maintenance still being a matter of debate, mental health professionals still encounter difficulties in diagnosing and identifying this mental disorder. The term Internet gaming refers only to video games. According to DSM-5, IGD can be diagnosed using a list of nine diagnostic criteria, requiring the occurrence of at least 5 in a period of 12 months. Some examples of these criteria are: withdrawal symptoms, constant preoccupation with video games, loss of interest in other activities, and lying/deceiving people or health professionals about the time spent playing, among others.

Five years later, in 2018, the World Health Organization (WHO) included the category of gaming disorder (GD) in the International Classification of Diseases (ICD-11) and classified it as behavioral addiction [16]. IGD is characterized by a pattern of continuous or recurrent gaming behavior manifested by only three symptoms: lack of control over gaming habits, prioritizing gaming over other interests and activities, and continuing to play games despite negative consequences. This pattern of gaming behavior may result in significant impairment in personal, social, educational, or occupational areas of functioning [16].

Mental health professionals, such as psychiatrists or psychologists, typically do the diagnosis of gaming disorder. They usually use diagnostic criteria outlined in the DSM-5 or the ICD-11 and require careful assessments by therapists. Also, a range of factors needs to be taken into account, such as severity, duration, and consequences of the gaming behavior, as well as any co-occurring mental health issues and family environment.

Currently, the diagnosis of this disorder uses a self-report tool as part of the assessment process, for that, gamers fill out a series of polls and surveys, answering questions regarding the gaming time spent, or the frequency by space of time, for example in Lyu publication [17]. Furthermore, the therapists use a combination of clinical interviews, assessment tools, and other sources of information, such as reports from family members or close friends, medical records, and behavioral observations [18].

A recent study identified and evaluated 32 screening tools that assess the phenomenon. They were based on the criteria outlined in the DSM-5 or the ICD-11 [19]. Many of these tools interpret the criteria freely and one of the challenges in the field is the lack of consensus on the criteria used for screening and diagnosing this condition. The diagnostic criteria for IGD have evolved over time and have been proposed by different organizations and researchers, leading to variability in the definitions used in various screening tools or diagnostic instruments. This can provoke big differences in prevalence rates, diagnostic accuracy, and comparability of research findings across studies [20].
In addition to the lack of consensus on assessing and measuring gaming behavior, the diagnosis of IGD can be challenging due to the subjective nature of some of the self-related reporting tools [21]. The main subjectivity problems come from the difficulty of remembering, by the patient, how many hours he spent during the week and how he felt during gaming.

In light of these challenges, researchers and clinicians continue to work towards establishing a consensus on the criteria used for screening and diagnosing IGD and develop a comprehensive standard assessment approach that considers the multidimensional nature of gaming behavior and its associated factors [19]. Nevertheless, researchers and mental health professionals are also exploring alternatives to those traditional assessment methods to improve the accuracy and validity of the diagnosis of IGD and to be used in conjunction with other clinical assessments and observations, as in the case of Lee investigation [22] which use cardiac rhythm.

With technological evolution, it has become increasingly easier to use data collection techniques (i.e., telemetry) on the events that take place during a game, generating large volumes of data that can be used for player analysis [2]. While telemetry data primarily focuses on player behavior and game performance, it can also be used to analyze the player from other perspectives, and help health professionals detect problematic player behaviors [23]. In addition, emotion data analysis tools that use audio and video information can be useful to identify the player’s emotional state during the matches [24] [25].

The project aims to address two key areas for improvement in the mental health field: the lack of consensus and the subjectivity in some diagnosis tools. One of our goals is to reduce bias between reality and self-reported responses. By exploring game data analysis and data on emotion we want to assist health professionals in detecting certain diagnostic symptoms, such as emotional fluctuations and time spent playing. Data from amateur and professional Counter-Strike players will be used to develop and test the proposed solution.

Joana Cardoso complements the research group [26], she is a Clinical Psychologist and Ph.D. candidate for the Universidade de Maia (Portugal), Departamento de Ciências Sociais e do Comportamento. Also, she is a member of the Portuguese Federation of Esports (FPEsports) [27].

Joana Cardoso, in terms of psychology and gaming behavior, is the group’s most experienced member, as her Ph.D. candidature is specialized in IGD. Furthermore, she works in a Portuguese clinic treating adolescents and young adults in their gaming use perturbations. Therefore, she was crucial in the entire process, firstly providing a large volume of related bibliography to Internet Gaming Disorder. Secondly, she described the current process of identification, diagnosis and treatment in the clinic where she works, which are the screening and intervention protocols defined in the DSM-5 manual. Then, after having regarded the available data, she oriented how the indicators and measurements developed (and described in Chapter 4) should be. Finally, most of the findings obtained during this project are based on the conclusions acquired by Joana Cardoso after the analysis of every single visualization, plot and table generated by the team.
1.2 Goals

The main goal of this project is to define metrics, based on telemetry and data on emotion, and to evaluate if they are relevant to help clarify and unify criteria for the diagnosis and treatment of IGD. The designed metrics aim to discern differences between gamers between them. These differences should be in terms of playtime, gaming intensity and emotional mood. In short, we try to answer these questions:

- How long time the player games?
- What is the intensity of these plays?
- How the gamer feels while he is playing?

By answering these questions, we will be able to accomplish the situation described in the motivation section, mainly to have an unbiased vision of the patient’s gaming behavior. Particularly, the proposed solution focused the developed metrics on the Counter-Strike game.

In short, the detailed objectives pursued look to find the data and extract value from it:

- **Find gaming open-source databases**: The data sources inspected, related to gaming telemetry, should include the login time, the initial and end datetime of every match, and the gamers involved in them. With these data, we could answer the first two questions related to telemetry, making possible the gaming behavior analysis and its evolution.

- **Extract emotions from gamer video selfie**: Regarding emotion data, the most relevant information we look for is the exact emotion the gamer feels while playing, second by second, to study the emotional fluctuations, answering the third question above. For that, we analyze video selfies of the gamer while he is playing. In addition, we want to analyze if these video emotion recognition tools are reliable.

- **Extract emotions from gamer speech**: As well as with video emotion, we want to extract the emotions by analyzing what the players say while playing. We also want to explore the reliability of this procedure.

- **Define useful metrics and KPIs**: One requirement for our solution is that these indicators should point to getting the most objective and readable information possible and looking for the most automatic and fastest updating.

- **Create visualizations for quick analysis**: These tables and charts are designed to be analyzed in parallel with the KPIs and in an efficient manner.

- **Validation of the KPIs and visualization**: A final objective of this proposal is to validate whether the designed metrics serve to differentiate individuals with abusive gaming behaviour from healthy ones.
1.3 Contributions

In this section, we explain the contributions of our project:

- Collection and analysis of the literature related to our project. Finding the benefits and drawbacks of the developed work in the area until now. As well as defining where our work could be useful and complement previous research.

- Define the main drawbacks and concerns of the current screening protocols, as well as viable ways to address them.

- Find multiple open gaming data sources as a basis for the accomplishment of the objectives. Discovering these data sources is crucial for the development of the project because they allow us to define the metrics and most relevant visualizations targeted. This data is stored in APIs (Application Programming Interface) and demo files.

- Put into practice the tools available for emotion recognition, mainly DeepFace for videos and Librosa for speech. These Python libraries allow us to process the gamer’s emotions while he is playing, throughout his speech and video-face recording during the game.

- Definition and design of the most relevant indicators and metrics, the tightest possible to the gamer’s reality. They inform about the number of hours played, per month and day, and about the most frequent feelings, among others.

- Definition of a set of visualizations that allow the therapist to understand the actual situation of the patient quickly.

- Use telemetry and data on emotion to answer most of the questions of the IGDS9-SF. This questionnaire is designed to diagnose IGD (more extensively explained below)

1.4 Structure of the document

In this section, we explain how the document is organized in order to facilitate its reading. Chapter 2 makes a series of briefing definitions of several concepts that are quite important to know before delve into the matter.

Chapter 3 introduces the reader to the related work. This section analyzes the most important approaches made until now regarding the Internet Gaming Disorder environment. For example, the different current screening tools and their characteristics. It also includes a subsection explaining the relationship between IGD and many other psychological comorbidities such as anxiety or depression.

Then Chapter 4, which is the most extensive. It describes in detail the proposed solution and introduces the Gaming data life cycle, which marks the guidelines and the steps to follow during the entire investigation. Starting from the data collection and all the data sources inspected, then
the description of the raw data obtained and how we transformed them into more complex and valuable metrics. Finally, we interpret the results and discuss how they can be useful, and at which moments, for the psychotherapists and their treatments.

Chapter 5 makes the summarization of all the processes in the section called Conclusions, where we also point out some of the benefits and drawbacks of our work. Then, for future investigations and continuations of our work, we open a series of research lines that could not be possible to tackle because a matter of time. For example to take a representative sample to make a statistical approach and to obtain richer conclusions.
Chapter 2

Background

This chapter clarifies several concepts that should be understood before going deeper into the subject. In the first section, we introduce what IGD exactly is according to the definitions of the DSM-5 and GD (Gaming Disorder) according to the ICD-11. Since we mainly focused on the IGD definition of DSM-5 over the entire investigation, we use the term IGD to encompass every definition related to gaming addiction (Including the ICD-11 Gaming Disorder definition). In the second section, we describe the possible risk factors that usually are related to IGD. In the third section, we introduce comorbidities, which are other mental disorders that may co-occur with IGD. The fourth section defines what emotions are according to different experts such as Paul Ekman. Section five makes a comparison between the two main behavioral disorders, IGD and Gambling Disorder. Section six describes the gaming motives which explain the reasons why people play video games. The seventh section explains the different game types and their characteristics. Finally, section eight introduces the Counter-Strike and its related gaming platforms.

2.1 Internet Gaming Disorder

Following the DSM-5 definition, Internet Gaming Disorder (IGD) can be diagnosed using a list of nine diagnostic criteria, requiring the occurrence of at least five in a period of at least twelve months [15]. Some of these items are:

1. Preoccupation with gaming
2. Withdrawal symptoms when gaming is taken away or not possible (sadness, anxiety, irritability)
3. Tolerance, the need to spend more time gaming to satisfy the urge
4. Inability to reduce playing, unsuccessful attempts to quit gaming
5. Giving up other activities, loss of interest in previously enjoyed activities due to gaming
6. Continuing to game despite problems
7. Deceiving family members or others about the amount of time spent on gaming
8. The use of gaming to relieve negative moods, such as guilt or hopelessness

9. Risk, having jeopardized or lost a job or relationship due to gaming

The World Health Organization (WHO) has its own definition of Gaming Disorder in the book “International Classification of Diseases 11th Revision: The Global Standard for Diagnostic Health Information” (ICD-11) [16]. According to it, gaming disorder occurs when the gamer presents a series of persistent or recurrent gaming behaviors, both online and offline. These patterns are manifested in 3 aspects:

- Impaired control over gaming (frequency or duration)
- Priority given to gaming over other activities or life interests
- Continuation of gaming despite negative consequences

The aftermath of these gaming behaviors results in distress or impairment in personal, family, social or occupational areas. Gamers may present symptoms continuously or by recurrent episodes [16].

Regarding the ICD-11 definition [16], individuals with gaming disorder may make numerous unsuccessful efforts to control their gaming behaviors, especially when imposed by others. Also, may increase the duration, frequency, or complexity of the games in order to maintain the levels of excitement. Individuals with gaming disorder, when cessation or reduction of gaming behavior, may experience dysphoria or disruptions in diet, sleep, exercise, or other healthy behaviors [16].

Gaming disorder diagnosis must be an ensemble of disorder features [16]. High gaming rates, long duration, or daily routine, individually, are not enough to determine IGD, neither having purposes of developing gaming skills and strategies, nor having an interest in social interaction is a sufficient basis to assign it, when alone.

Also, it is important to have into account some contexts such as age and gender, and cultural and peer-group norms, among others [16]. Their symptoms and consequences are not equal and for example, gaming disorder is more prevalent among males, and, adolescents or young adults. On the other hand, adolescents present anger control problems, emotional distress, and lower self-esteem, while adults are rather associated with anxiety and depressive symptoms [16].

Unlike ICD-11, the definition of IGD appears in the DSM-5 of APA, in the section of conditions for future study [15], with the intention of motivating discussion, but is not yet considered a disorder.

2.2 Risk factors in IGD

IGD can impact anyone, regardless of their personal, demographic, or professional situation. However, certain circumstances may increase the chances of being diagnosed with IGD. According to
Torres-Rodriguez et al. [18], these risk factors include a range of biological, personality, environmental, stress-related, and structural elements that make an individual more susceptible to IGD. These factors are crucial in determining one’s vulnerability to developing the disorder.

Biological aspects include predispositions to addiction and neurotransmitter deficits, while personality and vulnerability encompass emotional instability, low self-esteem, and inadequate coping mechanisms. Environmental factors like dysfunctional family dynamics and unsupportive school settings, along with stressors such as grief and major life changes, all contribute to the potential onset of IGD. Understanding these risk factors can aid in implementing effective prevention and intervention strategies to address the disorder’s prevalence.

Torres-Rodriguez et al. [18] agglomerate, in one of their reviews, the most relevant risk factors that can take part in the development of the IGD:

1. **Biological:**
   - Vulnerability to addictions
   - Deficits in neurotransmitters
   - Psychiatric comorbidity

2. **Personality and vulnerability:**
   - Immaturity and Emotional instability
   - Unconsolidated identity
   - Low self-esteem and indecision
   - Lack of self-control
   - Low resilience and frustration
   - High sensation search
   - Deficit of social skills

3. **Environmental:**
   - Family environment: Poor communication and affection, lack of supervision and family cohesion
   - School environment with low performance
   - Poor social environment

4. **Stress:**
   - Grief
   - Major crises
   - Drastic life changes
5. Structural

Numerous studies have examined how the risk factors contribute to the emergence and progression of IGD. Next, we expose some of those examples.

A study carried out by Rodriguez et al. [28], which looks for describe the PIPATIC program (Programa Individualizado Psicoterapéutico para la Adicción a las Tecnologías de la información y la comunicación), a Cognitive-Behavioural Treatment for 12- to 18-year-old adolescents with Internet Gaming Disorder (explained in section 4), reveals scholar and family problems in the pre-treatment phase in each of the 4 adolescent participants, such as bullying and exclusion by their peers, and family disorganization. Furthermore, 3 of them also presented comorbidities before the treatment, such as depression, social phobia, autism, and attention deficit hyperactivity disorder. After the Cognitive-Behavioural Treatment (CBT) all participants showed clinical improvement in the amount of time spent using video games and in the symptoms of IGD.

Furthermore, Celine’s study [29] affirms that non-problematic gamers are associated with families with better cohesion, while problematic gamers have poorer and more conflicting family relationships. In females, only, banning gaming activity by the parents is associated with higher levels of IGD. These findings confirm the strong relationship between parental attitudes and family structure on the occurrence of IGD. Thus, they conclude that prevention programs are needed to consider the importance of the family and parents’ influence on the problem and solution.

Melina et al. [30] associate IGD with parental rejection of gaming and CSE (Core Self Evaluation) scale which is a specific form of self-concept and personal appraisals that individuals fulfill regarding themselves, and it measures levels of self-esteem and neuroticism. Contrary to expectations, this investigation concludes that parental rejection positively correlates with IGD but not significantly. On the other hand, CSE is negatively correlated to IGD, hence, the lower the self-esteem, the higher the IGD Score.

Wartberg et al. [31] launched a longitudinal analysis with 1000 adolescents and they conclude that hyperactivity, inattention deficits and self-esteem problems seem to be important for IGD development. In addition, they found the first empirical evidence that IGD could contribute to the impairment of mental health in adolescents.

2.3 Comorbidities in IGD

Comorbidity refers to the presence of two or more medical or psychological conditions or disorders occurring simultaneously in an individual [32]. These conditions may co-occur and interact in various ways, potentially complicating each condition’s diagnosis, treatment, and prognosis. IGD, according to multiple studies [33, 34], is strongly correlated with other mental disorders. For that, in this section, firstly, we define several psychopathologies so related to IGD, because it is very important to understand them well. Next, we show some examples of those investigations that analyzed IGD comorbidities.

Regarding the DSM-5 definition, anxiety disorders encompass a group of disorders that have
as their central organizing subject the emotional state of fear, worry, or excessive apprehension. Anxiety disorders include panic phobia and social phobia among others. Specifically, social phobia causes significant distress when the individual is exposed to social situations or when there is a risk of being negatively evaluated by others. While panic phobia is characterized by recurrent and persistent concerns about having an attack.

Depression, following the DSM-5 definition, is similar to sadness, but more extreme and repeated over time. It takes more than days to disappear. It affects daily activities. Among its physical consequences are gain or loss of weight, pain, sleeping disruptions or lack of energy. Fortunately, depression can be addressed with therapy and antidepressants.

People who suffer Attention-deficit/hyperactivity disorder (ADHD) trouble staying focused, concentrated and organised, making realistic plans and thinking before acting. They may be constantly moving, fidgety, noisy and usually have problems adapting to changing situations. In the case of childhood, they may be defiant, socially inept and aggressive.

DSM-5 defines the Autism spectrum disorder (ASD) as a disorder characterized by difficulties with social interaction and communication. It is more typical among children. ASD includes autism, Asperger’s syndrome and childhood disintegrative disorder.

Obsessive-Compulsive Disorder (OCD) is a common anxiety disorder that involves intrusive thoughts or images (obsessions), accompanied by compulsive behaviors that help to decrease that anxiety.

Once we have contextualized the main mental disorders, we reference some studies concluding that IGD strongly correlates with other mental disorders. Especially on anxiety, depression, hyperactivity, and social phobia. Nevertheless, there is still controversy about the direction of these relationships, at least regarding depression.

Jeong et al. have found reciprocal relationships between IGD and depressive symptoms. They claim that understanding both disorders can assist in the prevention, treatment and remediation plans of these pathologies among children.

Analyzing the associations between IGD and its possible comorbidities, anxiety, depression and hyperactivity symptoms, they are significantly correlated with IGD in most part of the reviewed studies (around 90% of them). Also, social phobia is strongly related to IGD in the majority of the reviewed papers. 3 out of 4 reveal a significant correlation between both pathologies, IGD and social phobia. Hence, it would be recommended to study, in parallel, IGD with mental health issues.

Fazeli et al. conclude in one of their studies that depression and anxiety are strong mediators in the relationship between IGD and insomnia and quality of life. They conclude that families should pay special attention to the gaming time spent by children. Also, parents have to assist their children in coping with psychological distress.

There is also evidence that relates IGD to ADHD among children and youth. Adolescents with ADHD, apparently, have a greater risk for developing IGD. Anyway, it is not clear the proportion of ADHD children that will develop IGD. They also wonder about the direction of the
relationship between IGD and ADHD. For Muzwagi et al., due to the strong association between both pathologies, it is mandatory to consider ADHD as a risk factor for IGD and vice versa.

Murray et al. [38] also found an association between Autism and IGD. They launch an ANOVA comparing the average in IGD between individuals with and without ASD. The differences result significantly, ASD group is more prone to IGD.

Other studies relate psychopathology to the game genre and the game motives [39]. Low self-esteem and depressive symptoms are often analyzed to identify IGD. According to Billieux [39], self-esteem discriminates among the different game genres, especially between Casual (the highest self-esteem) and Action/Adventure (the lowest). Depression symptom levels are equal by game genre.

Individuals with low self-esteem and depression disorder have a high correlation with some of the game motives, such as escapism and avoidance coping [49]. Furthermore, the correlation face to IGD is significantly above zero (absolute value). However, only the depressive symptoms reveal a coefficient different from zero in the hierarchical regression when predicting IGD scores, among both genders.

Recent studies have also highlighted other no-mental health problems, such as sleep and diet disruptions [40]. This could be due to a lack of physical activity in players or because less physically fit individuals, such as those with obesity, tend to prefer less demanding activities like gaming. Analogously, people who spend extended periods gaming without physical activity may be at higher risk of sleep disorders [16]. In the particular Portuguese case, some studies [41] observed poor sleep quality, and psychological well-being scores were close to cut-off points. However, Portuguese gamers lead a healthy lifestyle, including regular physical activity and a healthy diet during gaming [41].

In conclusion, most of the bibliography related to IGD and psychological comorbidities agree on the importance of analyzing IGD together with the most common mental pathologies, because there may exist strong relationships [33, 36]. Even more, the direction of these relationships is unclear, sometimes suggesting that IGD could predict depression or anxiety, for example, [42].

2.4 Emotions

According to the Britannica Encyclopedia [43], an emotion is a complex experience that involves various aspects of consciousness, bodily sensations, and behavior, all of which reflect the personal significance or reaction to a particular thing, event, or situation. Emotions can influence how we perceive and respond to the world around us, shaping our thoughts, actions, and interactions with others.

Paul Ekman [44] defines emotions as natural responses shaped by our past experiences and evolutionary history. It involves an automatic evaluation of a situation that affects our well-being, leading to psychological changes and behavioral responses. Ekman defines six basic and primary emotions [45]: Happiness, sadness, anger, fear, surprise and disgust. The APA, in its dictionary of psychology [32] define them as:
• **Happiness** is an emotion of joy, gladness, satisfaction, and well-being.

• **Sadness** is an emotional state of unhappiness, ranging in intensity from mild to extreme and usually aroused by the loss of something that is highly valued.

• **Anger** is an emotion characterized by tension and hostility arising from frustration, real or imagined, injury by another, or perceived injustice.

• **Fear** is a basic intense emotion aroused by the detection of imminent threat, involving an immediate alarm reaction that mobilizes the organism by triggering a set of physiological changes.

• **Surprise** is an emotion typically resulting from the violation of an expectation or the detection of novelty in the environment.

• **Disgust** is a strong aversion, for example, to the taste, smell, or touch of something deemed revolting, or toward a person or behavior deemed morally repugnant.

According to Don Hockenbury and Sandra E. Hockenbury in their book "Discovering Psychology" [46], emotions involve three distinct components: a subjective experience, a physiological response, and a behavioral or expressive response.

In this document, we extract and analyze the emotions felt by several gamers. These emotional statuses are gathered from media files and represented in bar plots. On the other hand, several previous studies [47] gave a sequential order to the emotions, depending on their valence (from negativeness to positiveness) and on their arousal (from the least intense to the most intense), as we can see in the figure [48]. Our bar plot representations try to be faithful to that order, hence we sorted them as follows: Anger, fear, sadness, disgust, neutral, happiness and surprise.
Chapter 2. Background

2.5 IGD compared with gambling disorder

IGD is included in the last DSM version, as a behavioral disorder, similar to Gambling Disorder, as both disorders do not involve substance consumption. Comparing the implications and consequences of both disorders would be essential to understanding their unique characteristics, risk factors, and treatment approaches. Such a comparison could shed light on how excessive gambling and internet gaming behaviors affect individuals’ lives, mental health, and well-being, potentially guiding effective interventions and support strategies for those affected by these behavioral perturbations. While gaming activities’ consequences are associated with the disruption in diet and sleep [41], among others, Gambling Disorder is related to a high risk of bankruptcy [49]. Both conditions share family member relationship difficulties. Therefore, we have emphasized the most captivating methodologies, literature, and outcomes that are relevant to both conditions as discovered in the bibliography.

A Canadian study published in 2019 reveals that the majority of video gamers reported gambling in the past year (78.5%), and most past-year gamblers declared having played video games in the previous year (70.7%) [50]. The study found differences between problem video gamers and problem gamblers. Problem video gamers were generally younger, less impulsive, less likely to have a substance use disorder, and more likely to have depression. Although they shared some similar risk factors and characteristics, there was only a modest overlap between the two groups, with a small percentage being both problem gamblers and problem video gamers. The evidence suggests that while problem gaming and problem gambling have similarities, being heavily involved in one does not strongly predict involvement in the other [50].

Another investigation, from Spain, led by Mallorqui, [51] also compared IGD and Gambling Disorder individuals. Both clinical groups exhibited elevated psychopathological scores and less functional personality traits than a normative Spanish population. However, when comparing those with IGD to those with online Gambling Disorder, they appear to have several unique differences. Patients with IGD tended to be younger, more likely to be single and unemployed, and had a lower age of disorder onset. Additionally, they had lower somatization and depressive scores (contrary to the Canadian study), a lower prevalence of tobacco use, but higher scores for food addiction and a higher mean body mass index. Finally, they displayed lower novelty-seeking and persistence traits.

Another study carried out by Spanish researchers analyzes and compares the prevalence and relationship between gaming and gambling. They found a positive correlation between both disorders. Additionally, patients classified with IGD had poorer health-related quality of life [52].

A Korean investigation compares IGD and Gambling Disorder with Alcohol Use Disorder (AUD) and healthy individuals [53]. Among others, the questionnaire BIS-11 (Barratt Impulsiveness Scale version 11) was fulfilled by the participants in their investigation. The results show that IGD patients are more impulsive than healthy ones and those with Gambling Disorder. Individuals with AUD had similar behavior to IGDs. These findings not only contribute to understanding the categorical distinctions between individuals with IGD and Gambling Disorder but also shed light
on the dimensions of impulsivity and compulsivity that are involved in both conditions.

The design of multiple games has accessories and functionalities quite similar to gambling, such as the loot boxes. They are virtual items or in-game rewards in video games that can be purchased or earned by players. It is essentially a virtual mystery box containing random items, such as cosmetic skins, character outfits, weapons, or other in-game content. Players can use real money to buy loot boxes, and when opened, they receive a random selection of items. The contents of the loot box are usually undisclosed until it is opened, creating an element of chance and excitement. The use of loot boxes has generated controversy, as some consider it to be a form of gambling due to the uncertain nature of the items received and its potential impact on players, particularly young individuals. Furthermore, the European Parliament and the IMCO committee published a study analyzing loot boxes in online games around European countries. This document remarks that in countries like Belgium and the Netherlands, loot box purchases are considered gambling. Thus they are banned. Consequently, consumers in Belgium and the Netherlands do not have access to the full content of games compared with all other national EU markets where loot boxes are still legal.

Due to the relationship that could be between both conditions, several studies have analyzed Gambling Disorder treatments in IGD patients. Conclusions aim to the existent similarities between gambling and video gaming, such as both are repetitive activities, with intermittent reinforcement, decision-making activities, and risk-taking elements. Some people might assume that cognitive strategies used to treat problem gambling can also work for problematic video gaming. However, in this paper, they argue that cognitive approaches to gambling might not be useful for video gaming. These foci are based on erroneous beliefs about strategies and success in gambling (mainly about probabilities and randomness). Instead, they suggest focusing on other cognitive aspects related to gaming such as the over-valuing of gaming rewards, experiences and identities, the use of video gaming to maintain self-esteem, and video gaming for social status and recognition. This theoretical discussion serves as a starting point for developing more refined cognitive treatment methods for problematic video gaming.

In conclusion, there are several similarities between IGD and Gambling Disorder, such as they are both behavioral disorders. However, the differences are more significant. For example, gamers are younger and have substance use disorders less likely than gamblers, although they have higher scores for food addiction and a higher mean body mass index. Also, gamers are more impulsive than gamblers. Treatments that can be useful in gambling are not necessarily efficient in IGD. Some regulation has been established to protect gamers from gambling in some EU countries.

## 2.6 Gaming motives

Motives are sets of knowledge that reflect our emotional preferences, which are evident in our thoughts and ideas. They refer to specific aims that we desire or want to avoid, and they influence and energize our behavior.
The diverse dimensions of online gaming motives shed light on the multifaceted reasons driving gamers’ engagement. For some, it serves as an escape from their daily realities, offering a temporary respite from the stress or challenges they face [58]. Social interaction emerges as a significant motivator, fostering connections and friendships within virtual communities and fulfilling a need for belonging. The allure of competition fuels the thrill of mastering gameplay and achieving success. Skill and strategy development, satisfy the urge for personal growth and mastery [58]. Recreation serves as a form of leisure and entertainment. For certain individuals, gaming may act as a coping mechanism, confronting real-life problems or evading real-life issues [58]. Additionally, the appeal of living a fantastical life as someone else engages gamers in adventures offering them a thrilling form of imaginative exploration.

Analyzing these motives enables a better understanding of gamers’ needs and aspirations, informing game design and fostering a more engaging and fulfilling gaming environment. Demetrovics et al. [57] identified seven motivational dimensions for online gaming:

- Escapism from their realities
- Social interaction
- Competition
- Skill and strategy development
- Recreation
- Avoiding thinking in coping with real problems
- Fantasy (being someone else)

Manniko et al. [58] agree with most of the above mentioned motives. However, they define some others, such as experience immersion and achievement.

Recent investigations revealed that gaming motives discriminate the prevalence of IGD among gamers. In some, the motives are the strongest predictor [39]. In this research, Billieux et al. find escapism as particularly predictive, being higher in problematic gamers. Here, escapism is also highly and positively correlated with every motive, both psychopathologies (self-esteem and depression) and online time [39]. Furthermore, escapism motives are also associated with more elevated IGD scores regarding Manniko et al. [58].

In contrast, social motive does not always appear as a significant indicator of gaming perturbations. In the hierarchical regression trained by Billieux et al. [39] social motive shows a significant and positive coefficient to predict IGD. Also, according to Manniko et al. [58], the independent variable social motive appears as a positive and significant predictor of the linear regression for the variable problematic gaming score. However, in [59], despite of being analyzed, the social motive is not especially prominent in the three clusters classified as “Unregulated gamers”. In this approach, urgency, self-seeking, and self-esteem are more relevant. Achievement is also relevant when discriminating among IGD [59] [58].
Regarding avoidance coping and escapism, Melodia et al. [9] conclude that, if gaming alleviates and fulfills a real-life need, individuals may feel strong desires to go online life, and for a minority, it could lead to problematic consequences [9]. Clinical implications could not always be negative. Positive outcomes can arrive if avoiding off-line life makes the gamer feel better and implications are used for a short time. Analogously, if this mechanism becomes persistent and compromises other personal life activities can be considered problematic [9].

The desire to avoid in coping with real-life problems is also higher in gaming-disordered individuals [39], together with to be someone else for a while (fantasy).

At the beginning of their study, Manniko et al. [58], expected that motives would not be independent, so they developed a Principal Component Analysis (PCA) for dimensional reduction. Thus, categorized the motives into four families: (1) Experience-immersion, (2) entertainment and achievement, (3) social, and (4) escapism. Except for experience immersion, all the family motives contribute significantly to IGD prediction.

In conclusion, it is relevant to consider the gaming motives when analyzing IGD. Escapism and avoiding coping with real-life problems seem to be the most significant [39].

2.7 Gaming genres

The gaming world offers various genres to diverse player preferences. According to DSM-5 [15], considering the gaming genre is crucial when analyzing IGD due to variations in prevalence found across studies [60] [39]. Different game types offer unique mechanics and motivations that can influence players’ engagement and susceptibility to IGD. That was one of the main conclusions of Na et al. [60], who found differences in terms of average gaming time between the IGD-suspected gamers and healthy ones within the genre-specific groups. Hence, understanding genre-specific factors can help tailor interventions and preventive measures to address IGD and its associated challenges effectively. According to HP Tech [61], the main genres, by their relevance, are:

- Sandbox games grant players the freedom to explore and create within virtual realms.
- Real-time strategy (RTS) challenges strategic thinking as players control armies and resources.
- First Person Shooters (FPS) immerse players in intensive combat experiences.
- Multiplayer online battle arena (MOBA) games encourage teamwork and competition in fast-paced battles.
- Role-playing Games (MMORPG, Massively Multiplayer Online Role-Playing Games) enable players to embark on epic adventures, shaping their characters’ destinies.
- Simulation and sports games emulate real-life experiences and activities.
- Puzzlers and party games entertain with mental challenges.
• Action-adventure games merge thrilling action sequences with immersive storytelling.

• Survival and horror genres test players’ resilience in eerie environments.

• Platform games offer classic side-scrolling adventures boosting agility and precision.

Manniko et al. [58] expose examples for some of these game genres: For Role-Playing, Skyrim; For Action-Adventure, Tomb Rider; For FPS, Battlefield; For Platform, Sonic; For Sports, FIFA; For Puzzlers, Candy Crash.

Some investigations have taken into account the gaming genres in their tries to understand, predict and classify IGD individuals. Na et al. [60] in their research, conclude that MMORPG and FPS players are more prone to IGD than players of real-time strategy and sport games [60].

As well as with motives, Manniko et al. [58] did not expect genres to be independent of each other. Hence, they regrouped all the game genres into four features, again through a PCA. The first component encompassed Role-Playing games, including action-adventure, and FPS. The second group included simulation and sports. The third class integrated puzzles and mobile phone games. The fourth is fighting and music. Only the first group was a significant predictor of IGD [58].

Silvia Casale et al. [62] compare game genre with body dissociation. MOBA games show higher bodily dissociative experiences than MMORPG and FPS [62]. However, when the IGD score is greater enough the effect of the genre becomes indistinct. This finding suggests that players can disconnect from their bodies when the game engagement is high, no matter the game family.

Combining with motives, women only show differences when the motive is social and MOBA or MMORPG gamers [39]. This paper also suggests that, because of their characteristics, MOBA and MMORPG games, which imply many players gaming simultaneously, could reinforce the link between these genre games and social motives.

Gaming genres are also a crucial feature when investigating Internet Gaming Disorder. We have seen how action games, including FPS; and multiplayer game families, such as MOBA or MMORPG; are more prone to IGD than Puzzles or Adventure games [63].

2.8 Counter-Strike, Steam, FACEIT and Twitch

In order to accomplish our objectives, we chose the Counter-Strike (CS) game for the following reasons: it is a game with rich gameplay mechanics in which the players experiment with emotionally different conditions; the game is simple, even for a player having little familiarity with the genre; thus the game experience can be held as uniform as possible among participants involved in the research; and finally, because the psychologist Joana Cardoso as a therapist and member of the Portuguese Federation of Esports knows several CS players and consider an excellent choice to explore our ideas.

Counter-Strike is a game of the family FPS [64]. The gamer sees on the screen two hands holding a weapon, and the character is in a map or scenario in which there are other characters
(gamers or bots). The gamers are divided into two teams: Terrorists and Counter-Terrorists [65]. Terrorists win if kill every Counter-Terrorist member or plant the bomb and maintain it for 40 seconds. On the other hand, Counter-Terrorists win by killing all the Terrorists or rescuing the hostages.

There exist several game modes: competitive, casual, deathmatch, and war-game, among others. Competitive is comparable to CS played professionally online [65]. It consists of a thirty-round contest facing two teams of five members each. The match ends when one of the teams wins sixteen rounds [65].

There are numerous maps that can be distinguished by their size, if they have open areas or little hideouts and also, because of their background [64]. In some of them, like in the Office, the Counter-Terrorists have to release the hostages from an office without jeopardizing them. Besides, the Inferno, in which Terrorists have to bomb gas pipelines [64]. There are two more map types: Assassination and escape [66].

In total, there are 34 weapons in CS [67]. Among them, there are the knife, pistols, shotguns, rifles, and machine guns. Also, some other complements such as grenades, Molotov cocktails, and kevlar and helmet [67].

Players start with one hundred health points, and when they receive an impact they lose some of them, depending on where they receive the impact and the distance between shooter and victim [67]. When a player hits the head of an adversary it is called Headshot [68]. These shots are especially glorifying because they increase the 400% of the weapon’s base damage [68]. They are even counted in the API player stats [69]. Every time they kill some other player, they receive money for buying weapons, grenades or bulletproof vests in the next round [67].

Steam is an online platform developed by Valve [70] where the players can buy, play and discuss about gaming [71]. Initially, it was only for computers, nowadays also works on other console devices. This platform allows also you to buy a wide variety of games and complements, other than CS [71].

CS is the most popular game on Steam [72]. Also, Counter-Strike was one of the favorite FPS games in 2020 [73]. Its main attraction and popularity are due to its simplicity compared with other similar games [73].

FACEIT [74] is a platform that administers leagues for games such as Counter-Strike (CS), and it is linked to Steam [75]. In order to balance the matches, FACEIT has created a system of ten levels. If a team wins a match they win points, and the looser team loses them, like in a matchmaking system [75]. One of the main advantages of FACEIT over Steam is the Anti-Cheat software, which protects users from fraud and privacy.

Twitch [76] is a popular online platform where active users stream digital video broadcasts and passive users watch them. It was originally designed for video gaming, but due to popular expansion, nowadays, it also includes channels dedicated to artwork creations, music, sports and talk shows [77]. In this document, we use this platform to analyze professional gamers’ broadcasts, and particularly, to download the videos, from our participants, for later analysis and emotion
2.9 Summary

This section provided an overview of some of the key elements related to IGD, which are necessary to understand better the future sections, mainly the \[\text{Section A}\] and the \[\text{Section B}\]. The next section presents the main related works associated with our investigation, overviewing the current screening protocols and different approaches to address diagnosis and treatment for IGD.
Chapter 3

Related work

In the following sections, we describe some of the most relevant works related to our research by the different topics and approaches. Firstly, in section 1, we overview the most commonly used current screening protocols for diagnosis and treatment carried out. From sections 2 to 5, we explore different approaches to tackle IGD. They are based on physiological measures, emotion recognition and the use of technological tools, such as neuroimaging or telemetry, to identify gaming perturbations. While these alternative assessment methods are still in the early stages of development and require further research and validation, they offer promising avenues for improving the diagnosis and understanding of gaming disorder.

3.1 Screening tool questionnaires

There are several self-report questionnaires that have been commonly used in research and clinical practice to assess IGD. It is important to note that these questionnaires would only be used if the therapist suspects that the patient could suffer from IGD, not for any patient arriving at the clinic with symptoms. Some of the commonly used screening tools for gaming disorder include:

- Internet Gaming Disorder Scale—Short-Form The IGDS9-SF is a self-report questionnaire developed by Pontes and Griffiths (2015) \cite{78} to assess symptoms of IGD. It consists of nine items that assess various aspects of gaming behavior, such as preoccupation with gaming, withdrawal symptoms, tolerance, loss of control, and giving up other laser activities. Also, relating personal relationships and gaming activity. It was validated in several countries and languages, such as in Portuguese \cite{79} or Spanish \cite{80}.

- Problematic Online Gaming Questionnaire (POGQ): The POGQ is a self-report questionnaire developed by Demetrovics et al. \cite{81} to assess symptoms of problematic online gaming. It consists of 18 items that assess the six-dimension structure determined after the factor analysis carried out by the authors.

- Internet Gaming Disorder Scale (IGDS): The IGDS is a self-report questionnaire developed by Lemmens et al. (2009) \cite{82} to assess symptoms of IGD. It consists of 27 items that assess
different aspects of gaming behavior. It also has a short form consisting in 9 dichotomous items instead of the original 27.

- Gaming Addiction Scale for Adolescents (GASA): The GASA is a self-report questionnaire developed by Lemmens and Valkenburg [83] to assess symptoms of gaming addiction. It consists of 21 items to measure seven underlying criteria (i.e., salience, tolerance, mood modification, relapse, withdrawal, conflict, and problems).

- Other questionnaires are Video Game Addiction Test (VAT): The VAT is a self-report questionnaire developed by Van Rooij and Schoenmakers (2012) [84] to assess symptoms of video game addiction; and the Gaming Addiction Identification Test (GAIT) [85] A screening tool for gaming addiction in adolescents.

Particularly, we focused on the IGDS9-SF. It is a questionnaire developed by Pontes and Griffiths (2015) [86]. It reflects the nine clinical criteria for IGD defined in the DSM-5 [87].

Next, we present the nine items of the proposed questionnaire by Pontes and Griffiths.

1. Do you feel preoccupied with your gaming behaviour? (Some examples: Do you think about previous gaming activity or anticipate the next gaming sessions? Do you think gaming has become the dominant activity in your daily life?)

2. Do you feel more irritability, anxiety or even sadness when you try to either reduce or stop your gaming activity?

3. Do you feel the need to spend an increasing amount of time engaged in gaming in order to achieve satisfaction or pleasure?

4. Do you systematically fail when trying to control or cease your gaming activity?

5. Have you lost interest in previous hobbies and other entertainment activities as a result of your engagement with the game?

6. Have you continued your gaming activity despite knowing it was causing problems between you and other people?

7. Have you deceived any of your family members, therapist or others because of the amount of your gaming activity?

8. Do you play in order to temporarily escape or relieve a negative mood (e.g., helplessness, guilt, anxiety)?

9. Have you jeopardized or lost an important relationship, job or educational or career opportunity because of your gaming activity?
Chapter 3. Related work

This questionnaire quantifies how the patient feels according to the nine DSM-5 criteria for IGD. The answer to these questions uses a Likert scale from Never to Very Often. Each question gives points from 1 to 5, so the scale produces final scores between 9 and 45. However, there is still no consensus on the most appropriate cutoff score. A study in China [88] established it in 32. Besides, in an Italian context [89], the cutoff is lower, 21, and in a Brazilian research [90], risky gaming was found in 16 points and to be diagnosed the recommended threshold is 21.

3.2 Approaches based on physiological measures

As mentioned in the introduction, multiple IGD research lines use self-related tools during the diagnosis and treatment phases, with the drawbacks that it entails, such as subjectivity. In addition, questionnaires and interviews have limitations in the data collection process. For example, it can be difficult to report a player’s behavior at a specific moment or event during the game, or the players may feel uncomfortable when being watched or questioned, inhibiting their actual gaming experiences [91].

An alternative approach to user self-reported methods in post-game sessions is Psychophysiology [92]. These methods involve deriving psychological states from physiological measurements like eye tracking, facial recognition, skin temperature, electroencephalogram (EEG), electromyography (EMG), electrocardiogram (ECG), photoplethysmography (PPG), Heart Rate Variability (HRV), and galvanic skin response (GSR). These physiological metrics are frequently analyzed in neuroscience studies to recognize and understand physiological reactions.

Physiological signals are acquired using specialized sensors placed on the human body and connected to a computational system for processing. Various technologies are employed based on the type of signal being monitored. When choosing a device, it is crucial to consider its impact on the game flow, ensuring minimal interference with the user’s gaming experience [91].

There are many use cases for using psychophysiological measures. Most of them are related to understanding players’ behaviors, preferences, and experiences while gaming. There is a field called Game User Research (GUR) dedicated to those purposes [91]. It involves employing various research methods and techniques to gather data and insights from players to improve game design, usability, and overall player satisfaction.

Nevertheless, there are also many investigations that applied physiological data-gathering techniques to facilitate the diagnosis of IGD and follow up the treatment process.

In a Korean study [93], researchers utilized electroencephalography (EEG) to compare neural activity in healthy controls and individuals with Internet Gaming Disorder while watching repeated video game presentations wearing a head-mounted display. The IGD group exhibited higher absolute powers of Delta and Theta brainwave activities in regions like the prefrontal, central, and parieto-occipital areas. These findings imply potential disparities in brain activity between IGD and non-IGD individuals, offering valuable insights into the neurological underpinnings of IGD.

A literature review of electroencephalography studies has been carried out for researchers from the United States. They identified several studies relevant to their topic. Out of these studies, five
used EEG, three used Event Related Potential (ERP), and one investigated Low Late Potential (LLP) in IGD. The reviewed literature discarded those studies that were not focused on finding neurological biomarkers. Several relevant markers can be useful as IGD diagnosis features [92].

Based on the findings of this study, it is possible that increased resting-state slow-wave activity (specifically beta and theta waves) could serve as a useful neurological biomarker for identifying Internet Gaming Disorder. Additionally, two specific components of ERPs, known as N100 and P300, may also be indicative of IGD when their levels are decreased. ERPs are electrical responses in the brain that occur in response to a specific stimulus or event. Additionally, elevated late low potentials (LLPs) can be a specific trait marker of cue-induced cravings in IGD [92].

Analyzing if individuals experience cravings for gaming could be very interesting for the diagnosis and treatment of IGD. According to Korean researchers’ investigation, it is crucial [94]. During their study, they induced gaming cravings in adolescents with mild to severe IGD by showing them short video clips of gameplay from addictive games. During this craving state, various biosignals were recorded, including PPG, GSR, and electrooculogram measurements. By analyzing changes in these biosignals, the researchers used a support vector machine to classify each participant’s craving or non-craving states. This study demonstrated for the first time that the electrooculogram could serve as a useful biosignal marker for detecting gaming cravings.

Heart Rate Variability (HRV) has also been analyzed for detecting patterns in IGD patients. Specifically, two Korean universities have used it to measure if there is control dysfunction while gaming [22]. They hypothesized that individuals with IGD would show phasic suppression of vagally mediated HRV during gaming, reflecting the executive control dysfunction. Changes in HRV were associated with the IGD severity. They conclude that IGD young males showed an altered HRV response during gaming, suggesting executive control troubles [22]. The dynamics between executive control and reward-seeking may be imbalanced during gaming in IGD.

In short, there are many studies using physiological measurements to gain insights into the psychological state of patients. These studies often analyze human involuntary activity such as the heart rate, brain activity or sweat from the skin, which are beyond the control of gamers. Consequently, the information obtained through these methods is highly valuable as it is objective and free from subjectivity. On the other hand, despite the progress made, further research is necessary. The reviewed research consistently indicates the need to continue investigating IGD using physiological biomarkers [22] [92].

3.3 Approaches based on emotion recognition

Emotion recognition is the process of identifying and understanding humans’ emotional states or expressions. It involves perceiving and interpreting subtle cues to discern emotions such as happiness, sadness, anger, fear, or surprise. This skill is essential for effective social interaction, empathy, and communication. In the age of artificial intelligence and human-computer interaction, emotion recognition technologies are also becoming increasingly important, enabling machines to perceive and respond to human emotions, which has applications in fields like customer service,
mental health, and entertainment [95].

The identification of the human emotional state is often made through facial expressions, body language, voice tone, or physiological responses. For example in the automotive research led by Martin A. Tischler [95], where they use speech parameters, facial features and physiological data to classify the emotions of their individuals.

In the analysis of speech data, data mining and knowledge discovery techniques have been employed to extract valuable information from the speech data. Subsequently, this extracted information is utilized to construct two classifiers. The first one is designed for detecting valence, which relates to the emotional positivity or negativity of the speech, while the other classifier focuses on detecting arousal, which pertains to the level of excitement or energy in the speech. These classifiers serve to categorize and understand the emotional content conveyed through speech data [95]. At the Technische Universität München, they have developed robust algorithms that localize facial components from images or videos, including the eyes, chin, cheeks, and lips. These algorithms enable the interpretation of facial expressions [95]. Subsequently, a pre-trained model classifies this information to deduce the prevailing facial expression visible at any given moment. They specifically identify the six universal emotions as defined by Ekman [44].

In the same study, the Fraunhofer IGD Institute [96] designed an accessible wireless sensor system called EREC (Emotion RECognition) for gathering physiological parameters related to emotions. This system measures skin resistance, skin temperature, and heart rate. The EREC system comprises a glove equipped with sensors for skin temperature and skin resistance, as well as a wrist pocket housing the necessary sensing electronics. This user-friendly and convenient technology offers a valuable tool for monitoring and studying emotional responses through physiological measurements [95].

Noldus Information Technology is a company that provides software and hardware solutions for behavioral research [97]. They develop technologies to assist researchers in psychology, neuroscience and human factors. Recently, they have coded some facial actions, such as the inner brow raiser which contributes to the emotions of sadness, surprise, and fear; the Cheek raiser which aids in the happiness emotion recognition; and the Nose wrinkler which contributes to the emotion of disgust [97].

All of these advancements and discoveries are also applied to IGD in several investigations. A study in Bosnia-Herzegovina in combination with the UK [98] has analyzed metacognitions and emotion recognition to predict IGD in early adolescence. The hypothesis suggested that both metacognitions and a limited capacity to recognize negative emotions would be associated with IGD and would be strong predictors of IGD. They conclude that individuals who struggle to recognize negative emotions may turn to IGD as a quick reward-seeking method to relieve psychological distress. In this research, it is also analyzed the meta-worries, which could be defined as the concern about being worried. Meta-worries can be positive or negative depending on whether they are constructive or reflect a more distressing perspective. The presence of negative meta-worry was observed to be an independent predictor of withdrawal and conflict in the context of IGD. This
finding suggests that negative meta-worry might help elucidate the connection between IGD and its more severe manifestations, such as withdrawal (indicative of dependence) and conflict (indicative of behavioral impact on social interactions) [98]. Both withdrawal symptoms and impact on social relationships are two criteria of IGD according to DSM-5 [15].

Ruth L. Diaz led a research [99] that investigated whether individuals who played two or more hours of violent video games daily, as opposed to those who did not play video games, showed a distinct pattern in recognizing five facial emotions. Participants completed a facial recognition task. The study found that individuals who played violent video games recognized fearful faces with higher accuracy and faster response times than non-gamers. However, they showed reduced accuracy in recognizing disgusted faces. These findings suggest that desensitization to violence, continuous exposure to fear and anxiety during gameplay, and habituation to unpleasant stimuli might be potential mechanisms that explain these results.

The only video emotion recognition references related to IGD are about gamers’ ability to recognize emotions in images of others, not about researchers analyzing the emotions of the gamers themselves.

### 3.4 Approaches based on neuroimaging data with Machine Learning

Machine learning (ML) is a subfield of artificial intelligence that focuses on the development of algorithms and models capable of learning and making predictions or decisions based on data [100]. ML finds applications in diverse fields, from image and speech recognition to recommendation systems, autonomous vehicles, and finances, but in this document, we focus on medical diagnosis. Particularly, there is a systematic review regarding the applications of ML in addiction studies [101]. They analyzed 17 articles of which 14 were substance addiction-related, such as tobacco, cocaine, alcohol, opioids and multiple substances consumption. However, they also revised 3 non-substance addiction studies, regarding gambling and gaming (only 1). Most part of them used supervised learning. They conclude that ML methods are increasingly used in psychiatry for addiction disorder diagnosis and for informing medical decisions, especially supervised learning [101].

A study conducted by several universities in Korea [102] has trained a ML model for IGD classification. By combining clinical variables, EEG, and Positron Emission Tomography (PET) features obtained from different sources. Clinical variables were obtained from several tests such as Behavioral Inhibition System (BIS) and Behavioral Activation System Scales (BAS) and more. Then, EEG and PET features were obtained from individuals’ brain scanning. After testing different models, they presented a study proposing an IGD classification model by integrating features from each distinct source to enhance prediction accuracy. To optimize the combination of all these features and obtain an accuracy, sensitivity and specificity over 80%, they applied a Multiple Kernel-Support Vector Machine (MK-SVM). This methodology indicates practical applicability in real-world scenarios, and it supposes a robust prediction model for IGD [102]. However, the publishers recognize some limitations. For example, the individuals must have all these features
available. Furthermore, it is difficult to know how much one specific feature contributes to the prediction model.

Deep neural networks are also widely used in the identification and diagnosis of IGD, for example by Wang et al. [21]. This investigation suggests that most of the studies on IGD lack objectivity due to current screening scales and subjective judgments of doctors, resulting in diagnosis limitations. In this paper, a stop-signal task was used to measure inhibitory control in patients with Internet Gaming Disorder (IGD) using prefrontal functional near-infrared spectroscopy (fNIRS), a neuroimaging technique that measures changes in hemoglobin concentration in the prefrontal cortex of the brain. Participants were previously categorized into healthy and IGD groups based on their scores on the scale. After testing around 7 different Deep Learning (DL) and ML algorithms, they concluded that a Convolutional Neural Network (CNN) was the most accurate to classify the individuals. The results also showed that the use of CNN in fNIRS is a reliable method for identifying IGD patients.

Another investigation that analyzes images through DL was developed in China [103]. It combines several Neural Networks algorithms to enhance higher accuracy. Using EEG images the researchers create a real-time emotional state detection system for gamers. This system is based on a hybrid neural network called Convolutional Smooth Feedback Fuzzy Network (CSFFN). Specifically, CSFFN rationally combines a convolutional neural network (CNN), a fuzzy neural network (FNN), and a recurrent neural network (RNN). Experimental results show that this methodology has a high recognition accuracy and noise resistance in identifying four emotional states (happiness, sadness, superiority, and anger) [103].

During the last few years, there has been a rise in neuroimaging studies on young adults and adolescents with IGD. In this context, an Israeli investigation used brain imaging for detecting individuals with Internet Gaming Disorder [104]. They associated IGD with increased brain activity in areas responsible for craving, reward, emotions, loss of control, and sensory–motor processing. They discover some behavioral disruptions in brain areas, such as cognitive control, emotion control, salience network and goal-directed activity [104].

In short, another review [105] confirms some of the conclusions obtained above, that there exist multiple brain alterations in IGD individuals, while healthy subjects have not. For example, they suffer hyperactivation in the anterior and posterior cingulate cortices, and hypoactivation in the anterior Inferior Frontal Gyrus (IFG) in relation to hot-executive function. Moreover, IGD subjects showed reduced gray-matter volume [105]. They also agree in one aspect, neuroimaging techniques combined with Deep Learning can improve IGD identification and diagnosis, but further progress is needed.

### 3.5 Game telemetry analysis

The increasing complexity associated with video games has been accompanied by various technical improvements, which have made the collection of information from a given game session a much more simple, common, and reliable task [106]. Nowadays, large volumes of game data
are recorded daily, through game telemetry that uses instruments and sensors to collect real-time data in video games [107] including data related to the game, and events in which the players participated, and the actions they performed throughout the game.

This data can be used for various purposes: Analyze how players interact with the game and its maps and items, and analyze players’ performances. The interaction of players is very useful for game designers to understand the usability, playability, difficulty levels and gamers’ decision-making [108]. The developers can use these datasets to identify patterns in player behavior and preferences, which can help them create more engaging and satisfying experiences. On the other hand, it can also be used to analyze the performance of individual players and teams, based on statistics and analytics that can help coaches and players improve their performance [2, 109]. For example, telemetry data can be used to track player movement, shooting accuracy, reaction times, and other metrics that can help coaches identify areas for improvement. Similarly, data on team performance, such as win-loss records, can be analyzed to identify patterns and trends that can inform coaching strategies and player recruitment. In esports, telemetry data is especially valuable for analyzing the performance of players and teams, as it can provide insights into player behavior and decision-making in real-time. This can help coaches and analysts make quick adjustments to improve performance during tournaments and matches.

While telemetry data primarily focuses on player behavior and game performance, it can also be used to analyze the player from other perspectives, helping health professionals detect problematic player behaviors [23].

An example of related work using telemetry data is the research carried out by Zendle et al [110], which uses mobile gaming telemetry to find playtime differences between worldwide countries. Certainly finding significant conclusions by region. However, they do not address IGD.

On the other hand, Padman, Redma et al. also use gaming data in one of their studies for analyzing health issues [111]. Anyway, they focus on mobile app gaming for pediatric obesity, not for IGD.

The most interesting publication found is written by Parry [112]. It uses digital device usage telemetry (or gaming) and login information to validate self-reported tools in a systematic review. It analyzes the correlation between these assessment tools and the telemetry logged time. 106 studies were revised. Only one was related to gaming, published by Kahn [113], the rest of the papers analyzed were related to phone and social media usage. Findings indicate a low association between self-reports and social media telemetry data, suggesting that the current self-report could not be an accurate tool. Particularly in the unique gaming use case, on average, players underreport 1.26h per week, and despite having a significant correlation between reported and actual playing time, it is effectively small [113]. It confirms one of our hypotheses, these assessment tools are too subjective. The correlation between telemetry and the answer to the questionnaires is weak.

A study led by Andrews [114] supports the assumptions of Parry and Kahn. It analyzes three metrics of smartphone usage: Number of uses, Total duration, and duration length. Conclusions
were not forceful when comparing the self-reported usage and the mobile phone’s actual usage. So, they conclude that self-reported indicators should be taken with caution.

Furthermore, Thompson et al. [115] also use gaming telemetry in their studies to understand complex skill learning. One of their conclusions is that the importance of the variables used in their ML tools shifts depending on the gamers’ expertise level. In conclusion, different datasets are crucial for understanding real-world contexts, empowering the importance of gaming data.

Analyzing telemetry by country, we see that Portugal is not one of the most gaming countries by the amount of time spent on video games, as shown in Figure 3.1 [116]. Here we see the ratios that Portuguese (and worldwide) people devote to video games per week. 76% of the Portuguese population dedicates less than 5h or does not even play. Instead of the rest of the world, in which 42% of people spent more than 6 hours per week [116].

![Figure 3.1: Hours spent on playing video games in hours per week. Source Statista](image)

Previous research in France revealed differences in gaming time depending on the genre game [39]. These differences are demonstrated through an ANOVA, facing average gaming time by game genre. The strongest are between MMORPG (Massively Multiplayer Online Role-Playing Games), FPS, and MOBA positively against casual and other games. Online time is the same among the different game genres [39].

A Finnish study has also found differences in gaming use, by gender. The weekly amount of time spent is higher among men than women [58]. In contrast, this paper concludes that there is no discrimination between age classes.

From this section, we conclude two aspects. There are multiple investigations regarding the internet, mobile devices, computers, and social media uses. However, there is a lack of gaming telemetry analysis related to IGD [113]; on the other hand, the studies carried out on IT devices telemetry conclude that the correlation between self-reported tools and reality is weak or nonexistent [112, 114].

### 3.6 Summary

This chapter provided an overview of the most significant studies related to our investigation about IGD. It started highlighting the most common screening protocol, such as IGDS9-SF or POGQ. Then it explained several approaches that served as guidelines for our research, for example, methodologies to extract physiological measures or recognize emotions. Also, the chapter highlighted the use of neuroimaging and ML to assist in the diagnosis of IGD. Finally, since telemetry
gathering techniques are crucial in our investigation, we summarize some of the previous studies that served as guidance for us.
Chapter 4

Process of knowledge discovery through game data

This chapter introduces in the first section, how the intervention protocol is done nowadays in one particular Portuguese psychological clinic that traits patients with Internet Gaming Disorder symptoms. Then, in section 2, we define the proposed solution and we describe where, in the current process, our tools and visualizations can be used to help professionals in the diagnosis and treatment of IGD. Section 3 describes in detail all the steps followed during the investigation and is divided into 4 subsections:

- **Gaming data life cycle** explains the steps that a data science research should follow, serving us as guidelines for the entire investigation.

- **Data collection** subsection describes how the raw data gathering procedure is done, including gaming telemetry and data on emotions.

- The design and development of the **complex metrics** that will be useful in the patient disorder’s diagnosis and monitoring treatment.

- **Interpretation of the results** subsection compares the gaming behavior of the 4 participants using historical and recent data, as well as answering the IGDS9-SF questionnaire.

Finally, in section 4, we discuss the usefulness of our solutions.

4.1 Current intervention protocol

In this section, as an example, we describe a Cognitive-Behavioural program called PIPATIC [18]. It treats adolescents with excessive gaming behavior by addressing the main difficulties related to IGD diagnosis and treatment: risk factors and high comorbidity. We focus on this program because of its proven efficacy in the process of treatment of IGD [28,117].

The PIPATIC program was designed by Alejandra Torres-Rodriguez [28] and its goal is to offer specialized psychotherapy for adolescents with IGD symptoms together with other comorbid psychological disorders. It follows previous standards, such as those developed by Hansen et al.
Chapter 4. Process of knowledge discovery through game data

The PIPATIC program has a duration of 6 months. Therapy is divided into 22 weekly sessions of 45 minutes. The PIPATIC program must be guided by an expert clinical psychologist. Therapies are carried out individually, peer-to-peer [28]. The intervention between therapist and patient employs innovative techniques and resources commonly used in psychotherapy, such as empathy, acceptance, trust, paraphrasing, clarification, synthesis, confrontation, interpretation, feedback, promoting abilities and responsibility, and encouraging feelings of self-efficacy [28].

The program is divided into 6 modules [28]:

- **Psychoeducational:** Psychoeducation is an ongoing process within Cognitive-Behavioural Treatment (CBT) that provides patients with reasons, reflective capacities, and support to quit abusive gaming [119]. The therapist conducts motivational interviewing to teach self-monitoring strategies. Together, the therapist and the patient choose the entire therapy’s objectives. They should be realistic, for example, to limit gaming time, not necessarily to quit. The gamer learns and discovers as the opportunity arises, i.e., when it reads about IGD. The main goal of Psychoeducation is to help patients understand the treatment procedure, as well as their families, separately. This usually takes three sessions.

- **Standard IGD intervention:** This five-session module aims to reinforce stimulus control, learning appropriate coping responses, cognitive restructuring and problem-solving related to addiction. Also, the gamer learns to deal with temptation by being exposed to supervised gaming. The therapist encourages the patient to regain his or her pre-gaming hobbies. The gamer is also invited to do his professional duties, or homework, in a room other than where he games.

- **Intrapersonal:** The third module takes five sessions. It addresses emotional strategies such as identity, self-esteem, self-control, emotional intelligence, emotional coping skills and problem-solving. The therapist stimulates positive and proactive thoughts in the patient in addition to emotional regulation techniques and activities.

- **Interpersonal:** This module aims to improve relationships with other people. They work on verbal and nonverbal communication skills, assertiveness play-role activities and answering styles. The module takes two sessions.

- **Family:** The family and close friends’ environments are very important in the treatment. During three sessions they worked the active listening, family communication and specific activities related to showing affection.

- **Development of a new lifestyle:** During this phase, they analyze the balance between the previous and current situation of the patient, they develop a list of possible new activities and anticipate strategies to prevent risk situations and relapses.

Additionally, the PIPATIC program left the rest of the sessions floating for being included in the module that the therapist considers. In short, the psychotherapist and the patient meet every
Chapter 4. Process of knowledge discovery through game data

few days or weeks. Before each psychotherapy session, the patient completes several questionnaires, including IGDS9-SF. The therapist analyzes the answers before the session to understand the extent of gaming behavior from the last session, its impact on daily activities, and the patient’s emotional state. Then, the session starts. At the beginning of the treatment, the sessions are more regular, every week. As it progresses, sessions can become more spread over time.

The therapist should address a series of topics to increase the probability of success: normalization of gaming, the gaming industry aspects, the design of games, normalizing feelings related to change, and discussing how problems develop [119]. In the next paragraphs, we discuss these themes.

Gaming is a common leisure activity and one of the most frequent entertainment hobbies in industrialized countries [119]. For instance, in some countries in East Asia region is broadly accepted, especially in urbanized areas where there is a lack of leisure opportunities [119]. For the normalization of gaming, it is important to count on your close ones, such as family, and peers primarily known on the internet [119]. Another positive aspect to emphasize is that gaming is a relatively safe and healthy activity, especially compared to other physically dangerous activities (i.e., drugs). However, patients should always be aware of IGD risks [119].

Many scholar gamers defend gaming arguing that playing video games is an artistic expression [119]. Psychotherapists should inform gamers about the gaming industry because it is a very big business with millions of dollars in budgets [119]. The concept of “creative expression” should be confronted with ”making a profit”. To a large extent, games are designed to make gamers spend money on in-game purchases or buying updates to balance out those budgets [119]. For those reasons make understanding the gaming industry is clinically relevant.

In addition, games are also designed to engage players for as long as possible [119]. Games include lots of little goals and new items to achieve, making the gamer stand alert and engaged. As they reach end-game activities it appears new ones [119]. Some game genres like MOBA, often create social obligations or pressures to make the gamer play regularly. Some FPS games, like Doom, give extra hidden health points, to give the belief of being almost dead when the character is not, giving exciting near-miss experiences [119].

As therapy progress, it is expected that players minimize their playtime. That can create on the patient several difficult situations, such as anxiety, frustration or boredom [119]. It is also very important for the client to understand that these feelings are typical when stopping play. These negative thoughts should be reframed as a symptom that the client is mentally aware and preparing for the challenges ahead [119].

Finally, it is also important to remark on how IGD problems develop. For example, a reliance on gaming activities in early life can create false expectations about the nature of gaming [119]. Gaming-related problems can develop quickly and become chronic [119]. Also, there may be periods in which the client has achieved a playtime reduction, so it can relax. Then it should understand that any unexpected event can provoke a relapse and the problem can reemerge after an abstinence period [119].
As mentioned above, the studies about IGD have grown. However, the definition of evaluation treatments is still scarce [120], and the variability of criteria caused by the lack of consensus makes difficult the identification, diagnosis and treatment of IGD. On the other hand, Griffiths, Kuss and Pontes conclude that to determine if a treatment was effective it is necessary to analyze gaming time, IGD symptoms (withdrawal, tolerance...), performances at school or work, participation in leisure activities or hobbies, and interpersonal relationships [121].

4.2 Definition of the proposed solution

So far, we have examined various methods for diagnosing IGD, including treatment techniques. Most of the reviewed work in this document uses metrics and indicators based on interviews, questionnaires, and other analogical methodologies. However, therapists may encounter issues as answers related to gaming time, for example, that may not be entirely realistic due to subjective information provided by patients [113].

As the introduction states, this work aims to complement the diagnosis and therapy of IGD by utilizing gaming telemetry. It assesses the time of day, day of the week, and duration and frequency of gaming sessions to measure the intensity of the patient’s behavior evolution over days, weeks, months, and years.

Moreover, analyzing the state of mind can contribute relevant information for a proper diagnosis and therapy monitoring. Hence, we also gathered data on emotion from video files in which the gamer shows his face during the game, and we transformed it into interpretable features.

After the data collection, we developed a series of indicators and measurements to help therapists in the different stages of the treatment, such as in the first consultations, in which the clinician will be able to analyze the historical gaming behavior of the patient, hence making an initial diagnosis.

Also, before each session, it is very important for the psychologist to analyze the most recent evolution of the client’s gaming behavior. Therefore, we defined specific indicators that allow the therapist to review the gaming activity and the session intensity from the last consultation (some days or weeks ago), as well as the emotional fluctuations during the week.

In addition, we designed the plots and indicators to try to answer the nine questions of the IGDS9-SF. Analyzing, for example, the evolution of gaming time, we can determine if gaming has become the dominant activity in the daily life of the patient, or if it has given up other activities due to gaming, which is one of the nine diagnostic criteria of the DSM-5 [13].

Thereafter, the therapist will be able to find possible disagreements between the answer in the questionnaire and the objective reality known through telemetry, which could respond to other DSM-5 diagnostic criteria, having deceived family members or therapists regarding the amount of time spent on gaming. The patient’s loss of sense of time can explain the mismatch between questionnaire answers and telemetry data. This means that the gamer may be unaware of the time spent gaming.

Finally, together with psychologist Joana Cardoso, we will discuss the metrics and plots de-
signed to validate their usefulness. All these indicators development will be discussed in the section "Feature extraction and metrics design".

In conclusion, the proposed solution is based on the definition of those metrics and plots, and their usability and direct application on the main IGD diagnosis tool, the IGDS9-SF questionnaire.

4.3 Data workflow

As already mentioned, the main goal of this project is to use data analysis techniques on gaming data and data on emotion, to assist health professionals in detecting certain diagnostic criteria and following up the therapy programs. Therefore, in this section, we explain how we have developed the solution step by step. In subsection 1, we introduce the gaming data life cycle, where we describe the necessary steps a scientist must follow when working on a data-project related to gaming. This cycle sets the guide of the knowledge discovery process. For that, we have divided the solution following its steps, broadly: data gathering, metrics development, analysis and interpretation of the results, and discussion of the proposed solution.

In the second subsection, Data Collection, we introduce the different data sources explored to acquire the most accurate information available. In subsection 3, we present the raw data extracted directly from those sources and the design of more complex metrics. Finally, in subsection 4, we interpret the results obtained from the metrics and visualizations developed before.

As we will see next, we found several limitations in data acquisition and preprocessing, for example, due to the size of the data calls to databases. All the steps carried out to avoid these difficulties are also explained in this section as well as the useless or rejected pathways examined.

One requirement is to collect the data automatically, not asking the gamers directly. Therefore, to extract and analyze the data, the chosen programming language to work with during the entire project is Python, on its version 3.7 [122].

4.3.1 Gaming data life cycle

The gaming data life cycle refers to the stages involved in managing and utilizing gaming telemetry effectively for analysis and decision-making purposes [2]. Figure 4.1 summarizes these stages.

- Firstly, it is very important to define the attributes and the objectives of the discovery process in order to direct the data search [2]. It consists in the selection of the information to collect, and the definition of features, metrics, and Key Performance Indicators (KPI) for future analysis. KPIs are metrics specifically selected to analyze whether an objective is achieved. They need some historical context to reach conclusions [2]. In our case, for example, it is very useful to know the number of matches per month. All the defined attributes will be discussed in the sections below.

- Data acquisition is the process of gathering relevant data from various sources within a game environment. Ideally, this information is collected through a gaming telemetry system.
Chapter 4. Process of knowledge discovery through game data

This step could include capturing player interaction, in-game events, or match dates. It requires written code to collect and store the above defined variables in databases. Usually, the data is not collected only from one source, but from many. Actually, they could have very different formats and locations. In our case, we have an API, demofiles and data extracted from videos.

- **Data preprocessing** takes care of controlling data consistency. This can involve cleaning and filtering the data to remove errors, missing data, outliers, or irrelevant information [2]. This stage must be done before any further analysis. Also, this step is responsible for the data storage in databases or data files, locally or in the cloud. The data we collected did not require a cleaning process due to the good state of the main data sources.

- In the **metrics development** step, we define new and more complex features, such as the ratio kill-death. This is done by aggregating, summarizing, or deriving new variables from the raw data. One of the most relevant features we have defined in this step is the session, which agglomerates every match played in one sitting (with a pause, between one match and the next one, lower than 30 minutes). There is a section dedicated to the metrics development in this document.

- Next, **analysis and evaluation**, it focuses on extracting insights and patterns from the processed data. The goal is to gain a deeper understanding of player behavior. We can train a predictive model, accept or refactor a hypothesis or even generalize it into theory. The main techniques applied here are exploratory visualization approaches, statistical inference, and machine learning [2]. There is also a section dedicated to the data analysis discussion in this chapter.

- The insights derived from data analysis are often presented visually to aid understanding and communication. **Data visualization** techniques, such as charts, graphs, tables, and interactive dashboards, are used to represent the findings in a meaningful and accessible way. It is the most efficient way to expose conclusions and solutions to an audience with little prior knowledge [2]. Visualizations can help identify patterns, trends, and anomalies more easily. Even, in the data preprocessing step, to locate outliers or missing data.

- Related to the visualization, there are also the **reporting and knowledge deployment** steps. These final stages involve interpreting the results of the data analysis and deriving actionable insights. These outcomes usually are displayed in plots and tables deployed in those dynamic dashboards in order to be easily interpretable and understandable. This final step of the circle is often the beginning of a new one [2].

This project has completed every step learned from the gaming data life cycle, and used it as guidelines. In the next subsections, we detail, by topics/sources, how we applied the data acquisition and preprocessing stages, the definition of KPIs and metrics, and the analysis and interpretation of the results.
4.3.2 Data collection

This section corresponds to the Attribute Definition, Data Acquisition and Data preprocessing steps of the data life cycle. The attributes we want to gather are all related to telemetry data, for example, the list of matches played by the gamer. From here, we can extract very rich information that can be useful to determine how many matches he played in the last days and their duration. On the other hand, we transform multimedia information (video and speech) into interpretable features that represent the emotional state of the gamer, such as anger, happiness or surprise.

The structure of this section starts with the definition of the participants’ characteristics and how we get them. In the next subsections, we explain how we extracted data from the different environments: Gaming data, video emotion recognition data and speech emotion recognition data.

Participants

To evaluate the proposed solution, we selected four Portuguese male participants. The data collected from them is public and open source. However, they had to meet several requirements to fit our needs. Firstly, we needed two amateur and two professional gamers, which means that earn money through gaming. Also, we required, at least the professionals, to frequently stream videos on Twitch. These videos must show their faces frontally, well-illuminated, and not too far from the camera, which were the most discriminant factors. Additionally, all four candidates must also have FACEIT profiles and play CS using this application to enable telemetry data access through the FACEIT API.

We started searching in Portuguese Hubs Communities, particularly SAW Gaming HUB [123]. After multiple data source explorations and gamer profile analysis, we discarded the least inter-
testing gamers and, finally, we kept the four most informative. Unfortunately, we only found two
gamers with video requirements properly matched. Consequently, we have gaming data (from
FACEIT) for all four participants, but data on emotion (from video) only for two of them.

Gaming data

In this section, we explore all the possible data sources which could give us access to telemetry.
Mainly: Steam [124], HLTV [125] and FACEIT [69].

As stated above, Steam is the primary software subscribed by gamers to play. It serves as the
manager for Counter-Strike. To play CS, players must install this application and download the
game to their computer. So, our first source for telemetry data was Steam. Through its main API
and accessing through the python library steam [124]. However, among its features, this library
does not return information about the player or its statistics. It only gives access to the Steam
community and the Steam marketplace among others. Other libraries were inspected but their
documentations are insufficient or they are deprecated and make them useless (i.e., steamSpy)
[126].

One of the most appreciated data we looked for is the Steam login history [127] as it allows us
to determine the amount of time spent online, thus the gaming time. This feature is very important
at the psychological level, as we can see throughout all the literature [128]. However, this data
source had two issues: it requires permission from the player to access their login history, and we
want a solution that does not depend on this; also, it was not possible to access on a large scale,
but on a per-player basis. As a result, we chose not to use this method.

Next, we analyzed an international competitions manager platform, HLTV (Half-Life TV)
[125]. It provided very rich telemetry information, including player, team, match and map data.
However, this platform only gives access to information for the official competitions, so it is not
possible to access all the games of a player and then obtain the gaming time.

Also, we attempted to contact Valve, the company responsible for developing Steam, Counter-
Strike, and many other games, to inquire about the information we were seeking. We also tried
reaching out to various forums, such as the Steam community [129] and some other unofficial
platforms [130], but we were unsuccessful in our attempts.

Finally, we explored FACEIT [74], the linked platform to Steam through its user profile. It
allows the gamers to interact, create profiles and organize them like in a social network. It manages
a set of games, including CS. FACEIT provides, as well as HLTV, very detailed telemetry data
through Demofiles (file.dem). With the exception that FACEIT provides every match of the player.
These files contain the richest possible information. A demofile is similar to a JSON file, it is
divided into several levels. We segregate the most relevant information into 8 different tables
which are described below.

Firstly, it is important to clarify what is a frame.

• rounds: Every instance in this table summarizes the performance at the end of every round,
such as the starting and the ending datetimes, the winner faction, the end reason of the round,
and the score at the end of the round.

- **kills**: Each row represents one kill in the match, including the killer, the victim and the assistant of the kill, the weapon, the frame, the map, and the exact location of the killer and victim on the map (X, Y, Z coordinates), among others.

- **frames**: A frame is the smallest unit of time, and it is equivalent to 128 milliseconds. So, it allows tracking almost every single movement and the status of the gamer at every moment. The frame table informs about the general information of each frame, such as its frame moment and the living players by faction.

- **playerframes**: Each instance gives information about the player situation at each frame, so every 128 milliseconds 10 rows are created in the table (one per player). The following features are the most important: the player localization on the map, its health points and the number of kills and deaths in that round. Also, the weapon it is using and if it wears a helmet or a kevlar, the cash it has, and the latency ping, at every frame.

- **grenades & flashes**: In these two tables, each time a player throws a grenade or a flash bang, a row is created. Including the throwing player position on the map.

- **bombEvents**: It refers to every bomb action, such as planting, defusing or exploiting. As well as the player in action and the round.

- **damages**: This table stores damages data, so every time one player is harmed a row is added.

This information is valuable for deep analysis of gaming behavior, but Demofiles are computationally expensive due to their large size, around 150MB per match. Actually, FACEIT stores these files for only two months post-match. On the other hand, Demofiles give too detailed information for our purposes. At the moment, it is not necessary to analyze this type of information.

In addition, FACEIT provides free access to a database through an API [69]. This database stores the game data. It includes information about players, matches, and tournaments. Also, scores, rankings, and leaderboards. The statistics stored by match, so we know the final scores at the end of them. At the team and player level, the most relevant information delivered is the aggregated match performances, such as the kills and deaths of every gamer. We describe below the most important endpoints from which we have gathered the data.

- The gathering data workflow started with the player id search. For that, we need to know the nickname previously. The nickname is similar to the user name, so it uniquely identifies each player. Through it, we called the Search endpoint and we obtained the player id.

- Once we get the player id, we looked for the player’s general information, such as the region, the skill level by game, or its historic gaming performances. This data is stored in the Player endpoint.
• Furthermore, with the player id, from the Matches endpoint, we collected information about all the matches of a player: The match id, the starting and ending datatimes, and the team information including the teammates. Datatimes allow us to calculate the duration of matches, the day part and weekday of every match, the continued gaming time, and other player metrics. In addition, we gathered detailed match statistics by player. For example, the number of kills, deaths, assists, or kill-death ratio.

• The Hubs endpoint allows downloading data from every nickname enrolled in a community also known as hubs. At some time during the project, we found it relevant to gather data at a large scale to understand the general behavior of Portuguese CS players, and to develop the metrics we will discuss below. For that, we collected information on one of the most populated Hubs in Portugal: SAW GAMING HUB [123].

• Some other endpoints we worked with were Championships from where it is possible to collect all their matches; Teams for the team stats; and Game for the CS characteristics.

FACEIT is very adaptable to our requirements because we can organize tournaments and matches whenever we want and invite our gamers. Then, we will have their gaming time and performances. For accessing FACEIT data the user needs to sign up, and then an API key is provided which will be used as a bearer to authorize the access. Telemetry data collected from FACEIT API ranges from January’21 to December’22. The goal of having two-year historic data is because seasonality is a very important feature, hence we can analyze and justify a singular behavior in a specific period in the year, for example, an increasing or decreasing gaming time during holidays or exam period.

Video emotion recognition data

In this section, we explain all the data sources inspected to collect video data as well as every solution found to the numerous limitations we faced.

Regarding the video analysis, we have mostly inspected two Python libraries that analyze the video and return a series of characteristics of the images, in which must appear a face: DeepFace [24] and Face Emotion Recognition (FER) [131]. Finally, after several tests, we chose DeepFace, because it is more complete, and covers the requirements of this project, such as the emotion recognition from face expressions.

DeepFace contains a series of functionalities that describe the image. This description defines 7 emotional aspects: anger, fear, neutrality, sadness, disgust, happiness, and surprise. Also, it estimates the gender and age of the person in the image, as well as the race: Asian, White, Middle Eastern, Indian, Latino, and Black [24].

In the case of gender and age, it gives a single answer, instead of emotion and race. In these cases, it gives a proportion out of 1, for each of the 7 possible emotions and the 6 defined races [24]. The goal of our research is to recognize the emotions that the face is showing. Although, this library has other functionalities that, at the moment, won’t be explored.
DeepFace is a facial attribute analysis framework for Python. Still, actually, it is a hybrid framework that wraps other 6 state-of-the-art models (a.k.a. face detectors): OpenCV, SSD, Dlib, MTCNN, RetinaFace, and MediaPipe [24]. The developers of DeepFace are: Google, Facebook, and Keras, among others. In contrast, FER only uses OpenCV [131].

Firstly, to understand the functionalities and the dynamics of the two tested libraries, the image analysis experiments started with personal photos. Hence, we learned the age, gender, race and emotion outputs of the methods. As well as, understanding that DeepFace offers better outcomes than FER.

Secondly, short personal videos were passed through the Python function we coded. Here, we discover the huge range of possibilities offered by OpenCV [132].

Further analysis were carried out with YouTube public videos from CS professional players. Particularly, short videos, because they were computationally less expensive. In which we could see game images and the face of the gamer. We make several tests distinguishing gamers by gender, age and nationalities. Short YouTube videos resulted very suitable for choosing the best face detector because they were so realistic and so light.

According to the literature about face detectors, RetinaFace and MTCNN are more accurate. However, OpenCV and SSD are faster [24]. Anyway, we made several tests to find the most appropriate face detector for us. Firstly we discuss the characteristics of the most interesting models recommended by [24].

- **OpenCV** is an open-source library for computer vision. It is the most common and used library in face analysis field. It allows to access the camera of the user’s device to record videos, edit and transform them. Particularly, the most interesting option OpenCV provides for us is feature extraction for the emotion recognition step [132].

- **MTCNN** is the most complex face detector. It is composed of 3 CNN models: P-Net, R-Net and O-Net. Due to its structure, it is quite slower than others like OpenCV [133]. However, one of the best functionalities of MTCNN is face alignment. So, if the user’s concern is accuracy, it should use MTCNN.

- **RetinaFace** (RF) runs on a Tensorflow CNN. Hence it is quite accurate [134]. This model provides a method that outcomes the position of the face, which could be very useful, as we will see next. Also, it has the face alignment option.

To test the most appropriate face detector, we made some experiments. For that, we passed three short videos with similar duration, around 15 seconds, and we ran them several times. In figure 4.2, we show the average running times. We chose light files to speed up the simulations. MediaPipe [135], dlib [136] and ssd [137] face detectors were also tested with similar results to RetinaFace.

RetinaFace and OpenCV models are much faster than MTCNN. However, RetinaFace is more accurate, according to [24]. So the most suitable is RetinaFace. After choosing the best face
detector, we continue with more complex experiments. The goal at this stage was to make real
testing for understanding the scope of the data on emotion analysis. Hence, we looked for several
professional players who stream and film their faces during their matches. Twitch video-sharing
platform was the chosen multimedia source.

The videos we had access to contained several consecutive matches. Around 5 matches on
average and with a duration of 4-5 hours. Videos were downloaded using the Freemaker appli-
cation [138]. Freemaker platform allows the possibility of downloading the video by selecting
the quality in pixels (p), some options are 1080HP, 720p, 480p, 360p, and 240p. Depending on
the video quality the size of the downloaded file is bigger or smaller. Specifically, looking for a
balance between quality and computational efficiency, we transfer videos with a size of 420p. The
average weight of these files is 4GB.

The video sizes became a concern because analyzing these files was so computationally expen-
sive, besides the high volume of videos we analyzed. These limitations resulted in large running
times and memory errors. Several solutions were carried out to handle this situation:

- The first option we tested was to reduce the quality. So, downloading the videos in 360p
  or 240p. This solution was dismissed because faces were too undefined, and might be too
  imprecise.

- Secondly random sampling, there were selected 5 rounds from the complete match, reduc-
ing the duration of the videos from 40 minutes to 5, approximately. Therefore, extrapolating
  the conclusions to the entire film. This option was also rejected because we lost too much
  valuable information.

- As mentioned above, the RetinaFace face detector model has an option for face localization.
  To address memory errors we locate the face from the first image of the video and and
  analyze only this portion of the frames throughout the film. Nevertheless, we do not have a
  way to verify if this approach is accurate.

- Finally, we carried out the frame reduction. By default, the analyze method of DeepFace,
  which returns the emotions of the face, synthesizes 30 frames per second. So, 1 frame every
  33 milliseconds. The goal was to reduce the ratio to 1 frame per second. We consider that
  we don’t need much more level of detail.

For that, we coded a Python function to take into account only one frame per second. With
this solution, we lose some information, but we can economize the memory and computation
concerns, ensuring image quality.
Data collected from the videos to extract emotions range from 6th March 2023 to 19th March 2023. These data domains are different from telemetry data because it is not possible to know the actual date of the videos loaded more than a month ago. The Twitch platform only informs how many months ago the video was uploaded.

**Speech emotion recognition data**

McFee et al. [139] define a Speech Emotion Recognition (SER) system as a collection of methodologies that process and classify speech signals to detect emotions embedded in the audio records.

There are three main features that are extracted from the SER systems:

- Mel Frequency Cepstral Coefficient (MFCC)
- Chroma
- Mel

**Mel Frequency Cepstral Coefficient** represents the short-term power spectrum of a sound. In short, it identifies the audio and discards other stuff like noise [139]. The chroma is used for harmonic and melodic characteristics of music, meaningfully characterized pitches of music in 12 different categories. Finally, mel computes the mel spectrum, which is commonly used to represent audio signals, as it provides a rough model of human frequency perception [139].

Sound files can be transformed into a data set, in which every row may represent a variety of time-frequency units: seconds, frames, or samples. Also, we can set frequency representations: hertz, constant-Q basis index, Fourier basis index, Mel basis index, MIDI note number, or note in scientific pitch notation [139].

The most appropriate Python library found is **librosa** [25] [139]. It provides all the features that an audio analysis needs. Using these variables we can achieve the emotion of the gamer, the same as in video emotion: anger, fear, neutral, sad, disgust, happiness, and surprise.

The first stage for analyzing the emotions through the speech was to find a way of extracting the audio. So, using the Python library **moviepy** and its method **write-audiofile**, we extracted massively the audio from the videos of the professional players analyzed in the video emotion section.

**Librosa** library is not as evolved as **DeepFace**. **Librosa** returns the sound features mentioned above (MFCC, Chroma and Mel) but they have to be trained to determine the speaker’s feelings.

There exist on the internet several pre-trained models ready to use. We used one of them developed by Rocco Jay [140]. This classification model was trained using the voice of 24 actors. Every actor record 60 audio files simulating the 7 emotions. The accuracy of this model is quite good, more than 0.80. So, we made some experiments.

Using this pre-trained model has its limitations, it is important to remark before deciding if is suitable or not. The main concern for using the Rocco Jay model is that it is trained with audio in the English language. Our audios are in Portuguese. So, probably they may be inaccurate.
Secondly, the audio files we extracted from the videos have second voices from the teammates, so it is desirable to clean them up before classifying them as any emotion.

In conclusion, Speech emotion data won’t be used in our final solution because of the inconclusive results, and lack of solutions to face the problems raised.

4.3.3 Feature extraction and metrics design

In this section, we discuss the Metrics development stage of the data life cycle, and we provide an explanation of the raw data we collected from FACEIT and DeepFace video analysis. Additionally, we describe how we defined more complex features, like Session. Finally, we identify the KPIs that the therapist can use to gain knowledge from the data.

Gaming data

As previously mentioned, FACEIT has two sources of information: the API and Demofiles. The API offers general information like the match’s final score or date and time. On the other hand, Demofiles provide more detailed data for each match, such as the exact moment a player kills another. In appendix A.1 we present a table with all the features gathered from FACEIT API. The most significant are: `started_at` and `finished_at`, and the `match_id`. In appendix figure A.2 we summarize the most important variables we could collect from the FACEIT Demofiles: `health_points` and if player `is_alive`, from player-frame granularity; and `tick` and `attacker` from kill granularity.

Many of these raw data variables were aggregated and combined to create more intricate features that provided additional value. Below, we explain the definition of the most important:

- **A session** represents all the matches played consecutively, from when a player sits in front of his computer until he stops playing. Two matches are considered consecutive when the time between matches (time that elapses between the end of the last played match and the beginning of the next one) is lower than 30 minutes. The 30-minute gap was chosen because of personal gaming experience and the variable `time between matches` distribution (60th Percentile). A match only belongs to one session, but one session can agglomerate several matches. There could be several sessions within a day.

- **Gaming time** refers to the total duration of a gaming session, which includes the time between the first and last match.

- Based on the match date and time, we defined two new features: **day part** and **weekday**. **Day part** can take three values: morning (from 6 a.m. to 1 p.m.), afternoon (from 2 p.m. to 9 p.m.) and evening (from 10 p.m. to 5 a.m.). **Weekday** indicates if the match is played during the midweek (from Monday to Friday) or during the weekend (Saturday and Sunday).

- When analyzing matches, two additional characteristics that may be important to consider are the **competition type** and **game mode**. The competition type is typically divided into Matchmaking, where FACEIT matches your team with opponents of a similar skill level,
and Championship mode, which corresponds to a tournament. Game mode refers to the number of players on each team, such as 5 vs 5.

From these complex features, we define the indicators and KPIs that we show in 4.8, 4.15, and 4.16. The most relevant features extracted are the amount of time spent gaming each month (Avg played hours per month), the number of matches played per session (Avg matches per session), and the average duration of each gaming session (Avg hours per session).

**Video emotion data**

From video analysis data extraction, we collect the emotion at any time. That is the percentage of each emotion per second, as we can see in figure 4.3 which highlights the first 5 seconds of the raw data extracted from the video analysis, that is the ratio of each emotion felt per second, including: Angry, happiness, fear, sadness, disgust, surprise, and neutral. As stated above, we downloaded multiple video selfies of Participants 3 and 4. We downloaded the videos from the video platform Twitch.

<table>
<thead>
<tr>
<th>seconds</th>
<th>angry</th>
<th>disgust</th>
<th>fear</th>
<th>happy</th>
<th>sad</th>
<th>surprise</th>
<th>neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>48.64</td>
<td>0.01</td>
<td>5.64</td>
<td>0.4</td>
<td>40.72</td>
<td>0.03</td>
<td>4.56</td>
</tr>
<tr>
<td>1</td>
<td>27.01</td>
<td>0.33</td>
<td>67.52</td>
<td>0.05</td>
<td>1.2</td>
<td>3.65</td>
<td>0.24</td>
</tr>
<tr>
<td>2</td>
<td>10.98</td>
<td>0.01</td>
<td>88.88</td>
<td>0</td>
<td>0.11</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>54.56</td>
<td>0.02</td>
<td>5.63</td>
<td>0.03</td>
<td>34.3</td>
<td>0.01</td>
<td>5.44</td>
</tr>
<tr>
<td>4</td>
<td>99.4</td>
<td>0.11</td>
<td>0.31</td>
<td>0.04</td>
<td>0.11</td>
<td>0</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Figure 4.3: Emotion raw data example. Each row in the table represents the percentage time felt per emotion and second.

During the initial stage, we conducted a detailed analysis of emotions, scrutinizing each second and round. However, we encountered a challenge in this research pathway, as the evolution of emotions was highly variable and unstable, as we can see in appendix figure A.4. In simpler terms, analyzing and drawing conclusions for individuals was inconclusive.

To enhance readability and better understand the gamer’s emotions, a viable solution was to smooth the tendencies of each emotion. This was achieved by taking the median of the previous 5 seconds for every instance. As a result, it became easier to identify whether the gamer was happy or angry, among other emotions.

Once we know how the gamers felt during the entire round and match, we want to understand the reason why they switch their emotions during a game. Therefore, we joint telemetry from FACEIT demofiles with emotion obtained from the video analysis by second. The goal was to see how the player feels when killing, dying, or starting or ending the rounds. Our findings, illustrated in figure 4.4, revealed that Participant 4 felt angry when killing two opponents but fear and sadness at the end of a round. This association suggests that further investigation is needed in future research.
Figure 4.4: Percentage time felt, by Participant 4, during one round per emotion and second. Highlighted three events, two kills and one round end.

<table>
<thead>
<tr>
<th>event</th>
<th>angry</th>
<th>disgust</th>
<th>fear</th>
<th>happy</th>
<th>sad</th>
<th>surprise</th>
<th>neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>kill</td>
<td>13.54</td>
<td>0.05</td>
<td>11.23</td>
<td>2.14</td>
<td>8.92</td>
<td>0</td>
<td>0.63</td>
</tr>
<tr>
<td>death</td>
<td>22.56</td>
<td>1.28</td>
<td>17.47</td>
<td>2.39</td>
<td>27.86</td>
<td>0.04</td>
<td>2.78</td>
</tr>
<tr>
<td>round init</td>
<td>37.83</td>
<td>0.02</td>
<td>36.58</td>
<td>0.04</td>
<td>17.75</td>
<td>0.02</td>
<td>2.4</td>
</tr>
<tr>
<td>round end</td>
<td>0.07</td>
<td>0</td>
<td>32.04</td>
<td>2.65</td>
<td>46.9</td>
<td>0.06</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Figure 4.5: Percentual emotion average felt during each event.

This approach, including analysis of emotions by player, event, round and match, is quite complex and hard to carry out. However, one solution is to analyze the average emotion the gamer felt by events and match, as shown in figure 4.5. This table shows the average emotion when killing, dying, and starting and ending the rounds during one match.

After conducting numerous tests and trying different approaches, we concluded that determining the average emotion experienced throughout the match was the most effective method. This involved calculating the mean of each emotion, regardless of the specific event. As shown in Figure 4.6, we observed the emotions felt by Participant 3 during five separate matches in a session. Our findings revealed that the primary emotions were neutral and happy.

Figure 4.6 provides a comprehensive view of the overall emotions experienced by gamers throughout each match, disregarding any sudden emotional changes that may occur during gameplay. However, this method results in a loss of detailed information. On the other hand, it is crucial to understand whether external factors or events, unrelated to the game, influence the players’ emotions, such as household or professional environments. Therefore, we have created a chart
that shows the gamer’s emotional state at the start (first 5 minutes) and at the end (last 5 minutes) of the match, as illustrated in figure 4.7.

4.3.4 Results and questionnaire interpretation

In this section, we discuss the latter half of the data life cycle, which encompasses Analysis and Evaluation, Visualization, Reporting, and Knowledge deployment. We provide a detailed explanation of how our proposed solution can prove to be beneficial. We discuss how each developed metric and chart can assist therapists with their diagnosis and treatments. It is mainly divided into 3 subsections, the first one is oriented to the first stage of the treatment, where the psychotherapist starts knowing the patient, as explained in section 4.1. To achieve this, the therapist needs to understand the historical behavior of the gamer from different points of view. The second section, recent gaming behavior, involves monitoring the therapy and analyzing its effectiveness. Prior to each consultation, it is useful to briefly review the patient’s behavior in the past few days or since their last office visit, including analyzing gaming telemetry and data on emotion. This perspective complements the questionnaires filled out by the patient before each consultation. Lastly, in the third section, we answer the nine questions of the IGDS9-SF test, which assists in diagnosing IGD.

As this is just a proof of concept, we did not require a representative survey sample to make a statistically accurate approach. Therefore, we cannot apply the conclusions to a larger population. However, as stated in the introduction, our team includes psychologist Joana Cardoso, who provided a series of reasoning and arguments to properly interpret the results for the four participants.

According to the ICD-11 book of the WHO suggests, two patients with the same gaming
time won’t necessarily be equally diagnosed. It is also important to compare with the patient’s personal situation, for example, whether it is a student or worker.

**Historical gaming behaviour**

In accordance with the literature, it is crucial for therapists to have a clear understanding of their patient’s situation from the outset. This involves analyzing the amount of time the individuals spend gaming per day, week, and month, and determining in which day-part they typically engage in their activities. It is also important to assess whether there are any variations between weekend and weekday gaming habits.

In figure 4.8 we have the most relevant indicators that help understand the historical gaming behavior of the patient, such as the number of matches or the total hours played. Also, the intensity can be analyzed through the average hours per session (All the games played with less than 30 minutes of difference between them), or compared the gaming behavior by weekday and part of the day.

The analysis of these indicators reveals that in 2021 and 2022 Participants 1 and 2 primarily played in the evening while Participants 3 and 4 had a more balanced day-part ratio. Overall, the four participants played approximately 75% of their games during weekdays (midweek hours accounted for 72% of the total week time).

Regarding the average hours per session, Participant 4 looks like the most intensive player, this could be explained because it is professional. The other professional player, Participant 3, has doubled its playtime in 2022.
In figure 4.9, we see the total monthly hours played by Participant 2. Green bars show the total hours spent during the evening. Blue bars represent the total hours spent during the afternoon. As the morning hours played were minimal for all participants, they were not included in the graph. The data shown covers a 24-month period, from January 2021 to December 2022.

Participant 2 plays mostly in the evening, as seen also in the table above (4.8). Typically, it spends 30 or more hours playing each evening every month. However, it doesn’t play in August, possibly due to vacation time. The evolution of the average hours played per month is quite stable, except for the last months when it has reduced the total playtime.

Contrarily, Participant 3 also played during the afternoon. The monthly played hours have changed in the last 12 months, as shown in figure 4.10 and the Avg played hours per month metric in figure 4.8.
In Figure **4.11** we can observe the monthly hours played by Participant 3. The blue bars represent midweek hours, while the red bars show weekend hours. Here, again, we can see the evolution of Player 3. In 2021, except in March, it played around 27h per month on average (figure **4.8**). In 2022, as we mentioned above, it has more than doubled the time spent on gaming (67h). This increase in playtime could indicate both, a change in their gaming behaviour or a possible transition to professional gaming. Again, there is one month in which the player decreased gaming time, it would be worth analyzing why.
In order to understand if there exists some seasonality we have defined the chart represented in figure 4.12. It shows the total monthly hours. Purple bars represent the year 2021 and yellows the year 2022. Apparently, every June, Participant 3 stops playing for some reason. Almost every yellow bar is greater than purples indicating that in 2022 gaming time has significantly increased.

Analyzing the duration of gaming sessions is a reliable method of measuring a gamer’s intensity. Figure 4.13 provides an accurate overview of the gaming habits of Participant 2. The blue
bars indicate the number of monthly matches, while the green bars represent the number of gaming sessions played per month. If a gamer only plays one match per session, the blue and green bars will be of equal height. However, if the participant plays all the monthly matches in just a few sessions, the bars will be extremely far apart. Therefore, the greater the difference between the number of matches and sessions, the more intense the player’s gaming behavior.

When comparing the number of matches and sessions, and the blue and green bars between participants 2 and 4, we can infer that Participant 4 seems to be more engaged. In the discussion section, we will further compare the intensity plots between these two players.

Figure 4.13: Participant 2 monthly number of matches vs number of sessions.

Figure 4.14: Participant 4 monthly number of matches vs number of sessions.
Recent gaming behaviour

It is crucial to monitor patients on a weekly basis, alongside their initial diagnosis. To help the therapists, we have created specific indicators to track their gaming habits between consultations. These key performance indicators (KPIs) are meant to be presented to the therapist prior to each session, providing insight into the patient’s recent gaming behavior. The data presented in these tables and charts can help therapists determine the effectiveness of their treatments. The KPIs allow for the analysis of both gaming time and intensity during sessions.

The metrics that the therapist could view on their device are depicted in figure 4.15. These metrics comprise of three KPIs, namely the total number of hours spent in the last 7 days, and the percentage difference compared to the previous week and month. Participant 4 has played almost triple than Participant 1 in the last week. However, its gaming time evolution is more stable, since Participant 1 increased 52% in the last 7 days.

![Figure 4.15: KPIs related to the gaming time during the last 7 days, compared with the previous week and month.](image)

Additionally, by referring to figure 4.16, the therapist can evaluate the frequency and duration of the gamer’s computer sessions in the previous week (Total sessions last 7 days). The table also shows the percentage changes compared to the previous week and month.

It seems that Participant 1 has been exhibiting more relaxed behavior lately. This may be due to the fact that their sessions are shorter, and he played 7 times during the week. On the other hand, Participant 4 tends to play more than one session per day (9 in total), and their sessions are much longer on average.

![Figure 4.16: KPIs related to the intensity and duration of the sessions during the last 7 days, compared with the previous week and month.](image)
Similarly, figures 4.17 and 4.18 offer additional insights to the data presented in table 4.16. The vertical axis displays the number of matches (represented by blue bars) and sessions played (represented by orange bars) per day, spanning from today to 12 days prior. Just like in figures 4.14 and 4.15, a larger gap between the blue and orange bars indicates more intense gaming sessions.

From these graphics, it seems that Participant 4 has a more intense gaming behavior than Participant 1. In the past 13 days, Participant 4 played 9 matches three times, which implies more than 6 hours in a day (in 2 or 3 sessions) since each match takes 40 minutes (as shown in table 4.8). In contrast, Participant 1 played only 3 matches on most days, equivalent to 2 hours per day, sometimes in more than one session. Both players’ gaming behavior seems to be stable between the two observed weeks.

Figure 4.17: Participant 1 daily number of matches vs number of sessions, during the last 2 weeks.

Figure 4.18: Participant 4 daily number of matches vs number of sessions, during the last 2 weeks.
Recent emotional behaviour

In this section, we explain the metrics and plots created and to be explored by the therapist just before each consultation, regarding emotional fluctuations.

Figure 4.19 reveals how Participant 4 felt during 2 weeks, from 6th and 19th March 2023. It aggregates 13 matches played in 5 sessions. Most of the time he felt negative emotions. As we see in the table, he was angry, fearful and sad, 41%, 21% and 18% respectively. Next, it would be very interesting to understand the variation between matches and determine whether the emotions experienced are constantly present or it is only noticeable at certain moments.

Figure 4.19: Percentual average emotion felt by Participant 4 during the last 2 weeks.

In figure 4.20, we focus on match-by-match emotional behavior (Each of the bars with a significant distance between them indicates different sessions). Again, negative emotions, such as anger, fear and sadness, appear to be predominant during every match of Participant 4. However, there are some differences between sessions. In the first two (6 matches), fear, together with anger, is the most significant emotion felt. Besides, in the rest of the sessions, anger became even more dominant and sadness overtook fear in importance.

It is worth considering whether emotions change throughout a gaming match or if they’re influenced by external factors. Figure 4.21 can help us answer these questions. Just like in figure 4.7, we can observe the emotions experienced during the first and last 5 minutes of each match. Figure 4.21 represents the last 4 matches analyzed in figure 4.20. As previously mentioned, anger is the most common emotion observed, furthermore, it appears to increase as the match progresses. In the first 5 minutes, anger levels are lower than in the last 5 minutes. Conversely, sadness decreases toward the end of the match. Although happiness levels are low, they also seem to decrease at the end of some matches.
Figure 4.20: Percentual emotion average felt during 5 matches of Participant 3. Separated by sessions.

Figure 4.21: Percentual emotion average felt on the first and last 5 minutes of each match during the last session (4 matches). Participant 4.
Answering the IGDS9-SF

In this section, we discuss how our Key Performance Indicators (KPIs), metrics, and plots answer the nine items of the IGDS9-SF questionnaire. The 9 questions of this screening tool serve the psychotherapist to analyze the gaming behavior and diagnose IGD. The insights gained from this analysis can help in two ways: firstly, by providing automated answers to the questionnaire, and secondly, by measuring the patient’s subjective experience in cases where there are notable differences between answers and telemetry and video analysis indicators. Notice that answering the questionnaire through the indicators does not diagnose the patient, it is just an instrument to guide the therapist. Furthermore, these use cases proposed below just complement the traditional methods and serve to support the therapist in his analysis. Next, we discuss which metrics are suitable for each question.

**Question 1: Do you feel preoccupied with your gaming behavior?**

This question tries to identify whether the patient constantly thinks about gaming and defining tactics and strategies even when he is not actively playing, as well as if gaming has become a significant part of their daily routine.

It is not possible to answer this question because we cannot access the patient’s thoughts. Therefore, we don’t know if the gamer is preoccupied and constantly thinking about video games. The most we can guess is the amount of time spent on gaming activities. Then, the therapist may observe whether gaming is becoming a primary daily activity of the gamer if the last 7 days total hours has increased significantly. This indicator can be analyzed in figure 4.15.

**Question 2: Do you feel more irritability, anxiety or even sadness when you try to either reduce or stop your gaming activity?**

The purpose of Question 2 is to understand the patient’s emotional state when he stops gaming. It is typical to stop playing when the gamer feels satisfied or tired, but he may also experience negative feelings, such as anger, for example when someone else is asking him to stop, which could indicate abusive gaming.

The therapist can analyze the emotions at the end of each match or session to understand if the patient feels negative emotions when stopping or reducing gaming activity. If negative emotions have increased during one session, the therapist could conclude that the gamer feels "irritability, anxiety or even sadness" when he stops gaming. Figure 4.20 shows the average emotion per match, grouping matches by session. It allows us to compare the emotional status evolution over several consecutive matches. Also, figure 4.21 sheds light on this topic. It compares the emotions at the first and last 5 minutes of each match, so it allows us to conclude if negative emotions have increased during the match. Hence, if negative emotions are greater at the end of the match compared to the beginning, it could be an indicator of irritability.

Analyzing the number of daily gaming sessions is also important. This KPI can also indicate difficulty in stopping. If a gamer stops playing for just a few hours before starting again, it could
be a sign of abuse. Recent gaming behavior indicators, such as weekly playtime (figure 4.15) and weekly intensity (figure 4.16), can illustrate this perspective. For example, we can see in those figures that Participant 4 played 9 sessions in the last week, so more than one per day. This means that at least one day played two sessions. If this situation is repeated over time, it could aware the therapist about troubles in stopping playing.

**Question 3: Do you feel the need to spend an increasing amount of time engaged in gaming in order to achieve satisfaction or pleasure?**

Answering this question aims to determine if the patient has developed a tolerance to gaming, leading them to spend more time playing in order to feel satisfied or excited.

Also, to answer this item, we analyze the time spent playing games, particularly if the patient is playing more frequently and for longer sessions. By examining recent gaming behavior indicators, we analyze if the time spent on in-game activities has increased or changed in the last few days. For instance, refer to table 4.15 where we infer if total gaming time has increased regarding to the previous week or month. Furthermore, table 4.16 shows if the duration of the sessions is stable or has increased face to last week and month. Analogously, figure 4.18 informs graphically about the total number of hours and the intensity of the sessions. In short, if these indicators reveal an increasing tendency regarding gaming time and duration of the sessions, we conclude that the gamer needs to spend more time playing video games, so he is developing tolerance to gaming.

**Question 4: Do you systematically fail when trying to control or cease your gaming activity?**

There are two perspectives to analyze if the gamer is controlling or ceasing gaming activity: By analyzing the reduction of the total number of hours dedicated to gaming, during the last few days; or by analyzing if the patient plays more than one session per day.

The indicators that better fit these approaches are table 4.15 and table 4.16. They indicate the total hours in the last 7 days and the ratio of hours/session, also in the last week. Both have to be analyzed together. If gaming time and intensity grew during the last week or month, we could conclude that the gamer is not achieving to cease gaming activity. Analogously, it is advisable to graphically analyze figure 4.18. It shows the relationship between the number of daily matches and the number of daily sessions. If usually the patient plays more than one session could indicate trouble with stopping gaming activity. A gamer who plays every day or whose sessions are longer probably wouldn’t be controlling his gaming activity.

Also in a more general way, we analyze total time spent per month, by analyzing figure 4.11 or 4.9, which show the gaming time by midweek/weekend and by afternoon/evening. If each month’s total gaming time is not being reduced, we could infer that the gamer is failing in his ceasing gaming attemptings.
Question 5: Have you lost interest in previous hobbies and other entertainment activities as a result of your engagement with the game?

One of the criteria for diagnosing a patient with IGD is whether they have abandoned their other interests due to gaming. The approach we follow to answer this question is that paying attention to other duties or activities could be difficult if the gamer spends a lot of time playing.

If the amount of time spent has increased in recent months, it is possible that the patient has lost interest in other entertainment activities. To confirm or refute this hypothesis, we analyze the duration and intensity of the patient’s gaming activity over several months. The most appropriate figures to illustrate this are [4.10] and [4.11] which display the gaming time differentiating by day part (afternoon and evening) and by weekday (midweek or weekend). People usually enjoy their hobbies during the weekend or afternoons, so if gaming time predominates during these periods, the patient likely has lost interest in other entertainment activities. However, it is important for the therapist to know the patient working or studying habits, to better interpret and obtain conclusions when analyzing this question and their indicators.

Question 6: Have you continued your gaming activity despite knowing it was causing problems between you and other people?

This question answers DSM-5 criteria regarding the relationship between the gamer, and its gaming activity, with its family members and close ones. The indicators we developed only analyze gaming activity and the personal emotional status of the gamer. They cannot inform about the gamer's relationship with external people. Therefore, none of the KPIs shed light to answer this question.

Question 7: Have you deceived any of your family members, therapist or others because of the amount of your gaming activity?

This question analyzes how the gamer’s close ones feel in relation to the gamer’s activity, similar to Question 6. Hence, we could not develop any plot or indicator that might help psychologists answer this question.

Question 8: Do you play in order to temporarily escape or relieve a negative mood (e.g., helplessness, guilt, anxiety)?

This question aims to determine whether a patient uses gaming to escape from negative situations in their daily life. To do so, firstly we need to understand the emotional status of the gamer before the session starts and analyze how negative (or positive) he feels. Then, to compare with the end of the session, concluding the evolution of its feelings. The most appropriate figure to shed light on this topic is [4.20] which shows the average emotion by match and session. Hence, if positive emotions predominate at the beginning of the session, we can reject the relieving negative mood hypothesis. Otherwise, we could confirm that one possible gaming motive is to escape from negative thoughts. Furthermore, if the emotions get even worse over the course of the consecutive
matches, we conclude that gaming does not help to cope or escape the initial pessimistic scenario. In addition, the therapist can analyze the figure 4.21, which informs about emotional status at the beginning and the end of each match.

**Question 9: Have you jeopardized or lost an important relationship, job or educational or career opportunity because of your gaming activity?**

Similar to question five, this question aims to determine if gaming impacts off-line activities, in this case, regarding educational and professional occupations rather than hobbies.

We could assume that if gaming activity has increased or the intensity of the sessions is higher, it could be difficult for the patient to engage in any other activity, such as a job or educational courses. Figure 4.11 lets the analyst understand the gaming time tendency, particularly during weekends or midweek. In the context of this figure, we see that Participant 3 has significantly increased its gaming time during the last months, so he has probably jeopardized his educational or professional career. Another interesting perspective to analyze is the gaming time by part of the day. If the gamer usually plays during work time, it is also possible that he is negatively affecting his professional career. This situation can be analyzed through the figure 4.9 Once again to answer this question it is important to know the personal schedules of the patient and when he usually spends time on those jobs or educational careers.

### 4.4 Fitness for use & discussion

Although we faced some challenges during the project, our solution is beneficial and can be useful for the therapists. Also, it is aligned with the proposed goals. Our solution presents an unbiased perspective that therapists can use in their daily consultations. It offers insights into the gaming activity, such as the duration and intensity of the gaming sessions, as well as the gamer’s emotional status. We gather the metrics without any match intervention. This is important because players may feel uncomfortable when being watched, thus inhibiting their actual gaming experience [91].

There are two ways to approach the metrics we have designed. The first is based on the type of data collected and the underlying metrics, while the second is based on how useful they are for solving the problem at hand. On the one hand, data and metrics focused on hours of play, metrics designed to analyze the intensity of sessions, and data related to emotions. On the other hand, we have metrics with a historical view, designed for the beginning of the treatment; metrics with a short-term view, which serves to monitor the patient’s follow-up to the treatment; and finally, all these metrics can be used to answer the IGDS9-SF questionnaire and help to measure the patient’s actual situation and subjectivity. Anyway all of them are suitable to cover the motivation and goals of the project.

Throughout the project, and also reading the bibliography, we encountered numerous challenges in diagnosing Internet Gaming Disorder (IGD). One of the main difficulties was the lack of consensus and variability in interpretations [120]. Additionally, IGD presents two unique challenges compared to other behavioral disorders, such as gambling. In gaming, it is much harder to
identify individuals with disordered behavior because money is not directly involved. A gambler who is not addicted can quit quickly after losing money; however, those who are addicted continue to lose. In contrast, gaming has no money involved, so both addicted and healthy individuals can continue playing, making it harder to differentiate between them. Additionally, usually, professional players spend multiple weekly hours, making it even harder to distinguish between player profiles.

On the other hand, it is important to note that an increase in gaming time from one period to another does not automatically mean that the gamer has become addicted. As we mentioned earlier, a diagnosis of IGD requires multiple factors to be considered. Therefore, increasing gaming time may not always indicate addictive behavior but rather an advancement from amateur to professional. There are several criticisms of the diagnostic criteria that make the work of therapists more difficult.

The data were collected from the gaming activity of 4 participants. These were chosen for their specific characteristics in order to make comparisons between them and draw conclusions about their differences. The four participants’ behavior also helps validate the plots and metrics designed to solve the problems presented in the motivation section and the solutions proposed in the goal section. Additionally, players 1 and 2 were characterized as amateurs. In contrast, Players 3 and 4 are classified as professionals because they devote most part of their time to gaming, such as the case of Participant 3 who doubled his gaming time from one year to the next. Also because they earn money from donations during their broadcasts.

Next, for the validation of the developed indicators and charts, we discuss the main differences found in participants’ gaming behaviors after analyzing them.

Analyzing differences between Player 4 and the other 3, in figure 4.8, Participant 4 presents a much more engaged behavior. His playtime is higher, it has increased from 2021 to 2022, and his sessions are much longer. On the other hand, Player 3 has more than doubled his playtime. Regarding the KPIs related to emotions, the solution proposed also discriminates between Participants 4 and 3. Player 4 expresses mostly negative feelings when playing, mainly anger, fear and sadness, as we can see in figure 4.19. Meanwhile, Player 3 shows a much calmer mood, even happiness (see figure 4.6). Also, figure 4.21 shows how changes the state of mind of Player 4, he is more angry and sad, and happiness decreases at the end of the match compared with the beginning.

If we compare Player 4 against Player 2, we also see some significant differences. In the case of Participant 4 played 67 matches per month on average in 2021 and 91 in 2022, he increased 35%. Nevertheless, the number of matches per session remained stable, around 5. So, the intensity stayed equal (see figure 4.14). In contrast, regarding figures 4.8 and 4.13, Participant 2’s gaming time remained stable from 2021 to 2022, the number of matches per session is lower than Participant 4’s, it remained around 1.8, and his sessions take, on average, 1.4h (1h24m), less than half. Participant 1’s gaming activity also remained stable over the two years.

On balance, the proposed solution provides a comprehensive and diverse overview of patients’
gaming profiles and their historical evolution. However, it is fundamental to complement the analysis of these metrics and plots with the traditional assessment methods to assist health professionals in detecting certain diagnostic criteria and improving the accuracy and validity of the diagnosis. In addition, it is crucial to understand the meaning and definition of each indicator and what information they provide to draw accurate conclusions. The evidence is that some participants display signs of addiction according to some indicators, yet according to other metrics, their behavior is normal. It is especially remarkable the case of Participant 4. He has been engaged in long-lasting sessions, even more than one per day, and experiencing negative emotions while playing, indicating frustration. However, his level of intensity remained stable over the two years, indicating control. On the other hand, Participant 3 has increased his session intensity and doubled their yearly game time, but their mood has not been negative. Hence, we suggest that further analysis and investigations should be carried out.

4.5 Summary

The chapter Process of knowledge discovery through game data covered all the steps carried out during the entire investigation, from a review of the current intervention protocol and an extensive definition of the proposed solution to the interpretation of the results. Passing throughout the participants’ selection process, the data workflow, which included how telemetry data was collected and emotions were gathered from videos, and the feature extraction and metrics development. Ending with a discussion of the results.
Chapter 5

Conclusion & forthcoming work

This chapter summarizes the most relevant topics we addressed throughout the document, remarking on some of the strengths and weaknesses from which we extract the main pathways for future work.

5.1 Conclusion

In the present document, we described all the processes, step by step, carried out to accomplish the goals set at the outset. These objectives included finding open source data for gaming telemetry and video emotion analysis, defining complex metrics, and creating visualizations to aid in identifying and diagnosing IGD as well as monitoring the progress of associated therapies.

It started organizing the document by identifying the motivation and objectives, which is crucial in ensuring clear guidelines and efficient project completion. To enhance comprehension, in the background chapter, we included several key concepts that should be understood before delving deeper into the subject. This includes defining Internet Gaming Disorder and identifying the risk factors that can lead to it.

Prior to commencing our work, we thoroughly examined a vast array of publications related to IGD. Our project has been built upon the foundations laid by these works. We have identified the most pertinent studies and made a comparison between various approaches to diagnose and treat the IGD. Namely, approaches based on: physiological measures, emotion recognition, neuroimaging data with ML and gaming telemetry analysis. This section is quite informative as it highlights the advantages and disadvantages of the existing screening and intervention tools.

Thereafter, our contribution begins by explaining and executing our proposal. Firstly, obtaining the data and processing it. Then extracting value according to the guidelines set by the data life cycle and Psychologist Joana Cardoso’s knowledge. This involves defining KPIs and creating visualization charts, which is the ultimate value we want to extract for the usefulness of our solution.

Despite being satisfied with our work, we recognize that there is still room for improvement. We have encountered several obstacles, for example, the challenge of obtaining free, accessible, and accurate data. Despite examining multiple data sources, they were all rejected due to various
shortcomings, including insufficient historical data. Additionally, we faced difficulties when processing large video files, particularly disc space and memory problems. However, we were able to find solutions to these issues. Unfortunately, we could not improve several aspects of our work due to these complications, which we outline in the following section for future work.

5.2 Forthcoming work

Our proposed solution aims to address the issues presented in the current diagnosis and treatment tools for Internet Gaming Disorder. This includes the absence of objective information regarding the patient’s emotions while gaming, and the true duration of their video game use. The solutions we gave go through finding open data sources and data analysis of gaming telemetry and data on emotion. To achieve that, we designed a series of metrics and charts to give the therapist a better point of view of the patient’s situation.

Although the proposed solution is very useful and shows great potential, it is still in its initial stages. We acknowledge several areas where improvements can be made, namely to obtain a proper sample to develop a significant statistical approach, to gather data for all of the available games, to train ML speech and video emotion recognition models and to develop a dashboard application for therapists to use in their daily work.

As part of this project, we aimed to create metrics and visualizations that would provide psychotherapists with multiple perspectives on their patients. We selected four participants with diverse characteristics including varying gaming behaviors, engagement levels, and amateur/professional profiles. However, it is important to note that this sample size is not suitable for any statistical approach. We can not conclude which combination of features is determinant for classifying an individual as disordered, nor can we establish any thresholds based on monthly gaming hours or the number of sessions. Instead, we have defined the fundamental indicators that future research should analyze.

So, the next step could be to collect the correct sample of individuals and measure their gaming behavior over several months. This will help identify which features are statistically significant for determining IGD and if there are any notable differences between healthy and disordered gamers.

The proposed metrics and plots exposed above are based strongly on gaming time. That is to analyze how long time the gaming sessions take and their intensity. Nevertheless, it is important to note that these KPIs and conclusions were derived solely from telemetry data collected from the Counter-Strike game. Therefore, our understanding of the patient’s overall gaming habits is limited to their activity within this specific game. Consequently, if the patient doesn’t play Counter-Strike on weekends, it may suggest that gaming isn’t a daily activity for him and isn’t interfering with their real-life commitments. However, it is possible that he is playing other games.

Our proposal for future work involves opening telemetry data to all available games. By analyzing the developed plots and indicators, taking into account all the gaming activity and using a proper statistical approach, we can gain a more realistic understanding of the situation.

Analyzing speech data was one of the initial goals, and we worked on it thoroughly. However,
we encountered a challenging situation to overcome: There were no useful pre-trained Machine Learning models available to classify speech into emotions freely accessible on the internet, as it was for video emotion analysis. The available classification models were trained in English, not in Portuguese such as our audio files, and training data was not in the correct environment, not in a gaming streaming. Therefore, the accuracy of these models would likely be compromised.

The solution requires training a model exclusively for our purposes. The best way to proceed is to train our own model. For that, we should trim a sample of audio files and label them based on the emotions they convey. Using the labeled files, we will train the model to classify participants’ emotions accurately. However, due to time constraints, we were unable to pursue this method.

FACEIT API is open-source, and it is currently used in several applications. Therefore these software make calls to obtain the data and update their indicators. However, these applications are completely oriented to the gamers’ performances, like trackergg [141]. To our knowledge, there is no application using gaming telemetry for help in IGD diagnosis and monitoring.

Hence, in order to have a completely practical solution for the psychotherapist, it would be necessary to develop a dashboard application with access to the FACEIT API to regularly update the data. The goal is to provide a tool for the therapist to integrate the metrics and plots defined during this project including the upgrades proposed in the future work chapter, such as having data from all the available games.
Bibliography


Appendix A

Appendix

<table>
<thead>
<tr>
<th>Feature</th>
<th>Granularity</th>
<th>Description</th>
<th>Comments</th>
<th>endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>player_id</td>
<td>player</td>
<td>player identifier</td>
<td></td>
<td>players/player_id</td>
</tr>
<tr>
<td>player_nickname</td>
<td>player</td>
<td>player nickname</td>
<td></td>
<td>players/player_nickname</td>
</tr>
<tr>
<td>country</td>
<td>player</td>
<td>country code (PT)</td>
<td></td>
<td>players/player_id</td>
</tr>
<tr>
<td>skill_level</td>
<td>player</td>
<td>from 1 to 10</td>
<td></td>
<td>players/player_id</td>
</tr>
<tr>
<td>facukt_aka</td>
<td>player</td>
<td>integer points to balance teams</td>
<td></td>
<td>players/player_id</td>
</tr>
<tr>
<td>friends_db</td>
<td>player</td>
<td>We can count the number of friends</td>
<td></td>
<td>players/player_id</td>
</tr>
<tr>
<td>membership</td>
<td>player</td>
<td>premium, free or free</td>
<td></td>
<td>players/player_id</td>
</tr>
<tr>
<td>infections</td>
<td>player</td>
<td>integer</td>
<td></td>
<td>players/player_id</td>
</tr>
<tr>
<td>verified</td>
<td>player</td>
<td>true or false</td>
<td></td>
<td>search_players</td>
</tr>
<tr>
<td>status</td>
<td>player</td>
<td>active or available</td>
<td></td>
<td>search_players</td>
</tr>
<tr>
<td>total_number_won_matches</td>
<td>player</td>
<td>integer Lifetime total number of matches played</td>
<td></td>
<td>players/player_id/games/game_id</td>
</tr>
<tr>
<td>longest_win_streak</td>
<td>player</td>
<td>integer Lifetime statistics</td>
<td></td>
<td>players/player_id/games/game_id</td>
</tr>
<tr>
<td>wins</td>
<td>player</td>
<td>integer Lifetime number of won matches</td>
<td></td>
<td>players/player_id/games/game_id</td>
</tr>
<tr>
<td>K_D_Ratio</td>
<td>player</td>
<td>integer Lifetime Kill/Death ratio</td>
<td></td>
<td>players/player_id/games/game_id</td>
</tr>
<tr>
<td>Average_K_D_Ratio</td>
<td>player</td>
<td>integer Lifetime statistics</td>
<td></td>
<td>players/player_id/games/game_id</td>
</tr>
<tr>
<td>recent_results</td>
<td>player</td>
<td>float array [win, lose] last 5 results, 1 win : 0 lose</td>
<td></td>
<td>players/player_id/games/game_id</td>
</tr>
<tr>
<td>total_headshots</td>
<td>player</td>
<td>integer Lifetime statistics</td>
<td></td>
<td>players/player_id/games/game_id</td>
</tr>
<tr>
<td>win_rate</td>
<td>player</td>
<td>percentage Lifetime statistics</td>
<td></td>
<td>players/player_id/games/game_id</td>
</tr>
<tr>
<td>average_headshots %</td>
<td>player</td>
<td>percentage Lifetime statistics</td>
<td></td>
<td>players/player_id/games/game_id</td>
</tr>
<tr>
<td>match_id</td>
<td>match</td>
<td>match identifier</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>competition_type</td>
<td>match</td>
<td>string matchmaking: championship</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>competition_name</td>
<td>match</td>
<td>string</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>started_at</td>
<td>match</td>
<td>start time</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>finished_at</td>
<td>match</td>
<td>end time</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>winner</td>
<td>match</td>
<td>string</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>time</td>
<td>match</td>
<td>formatted or fraction</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>result_faction</td>
<td>match</td>
<td>integer</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>feature</td>
<td>Granularity</td>
<td>Description</td>
<td>Comments</td>
<td>endpoint</td>
</tr>
<tr>
<td>round</td>
<td>match</td>
<td>integer</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>best_of</td>
<td>match</td>
<td>integer</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>game_mode</td>
<td>match</td>
<td>string</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>round_date</td>
<td>match</td>
<td>integer</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>round</td>
<td>match</td>
<td>integer</td>
<td>total number of rounds</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>winner</td>
<td>match</td>
<td>string</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>score</td>
<td>match</td>
<td>string</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>map</td>
<td>match</td>
<td>string</td>
<td>Name of the map ( inferno, anomaly,...)</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>team_id</td>
<td>match</td>
<td>match Identifier</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>team_win</td>
<td>match</td>
<td>matchIdentifier</td>
<td>true or false if the team won the round</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>final_score</td>
<td>match</td>
<td>matchHeader</td>
<td>final score of the round</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>final_kills</td>
<td>match</td>
<td>matchHeader</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>team_headshots</td>
<td>match</td>
<td>matchHeader</td>
<td>teamHeadshots percentage</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>player_id</td>
<td>match</td>
<td>matchIdentifier</td>
<td></td>
<td>matches/match_id</td>
</tr>
<tr>
<td>kills</td>
<td>match</td>
<td>integer</td>
<td>Number of kills</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>triple_kills</td>
<td>match</td>
<td>integer</td>
<td>Number of times the player kill 3 enemies without being killed</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>quadra_kills</td>
<td>match</td>
<td>integer</td>
<td>Number of times the player kill 4 enemies without being killed</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>perma_kills</td>
<td>match</td>
<td>integer</td>
<td>Number of times the player kill 5 enemies without being killed</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>assists</td>
<td>match</td>
<td>integer</td>
<td>Number of times the player hearts without killing the enemy and a teammate does</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>headshots</td>
<td>match</td>
<td>integer</td>
<td>Number of headshots</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>headshots %</td>
<td>match</td>
<td>percentage</td>
<td>Percentage against total of kills</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>deaths</td>
<td>match</td>
<td>integer</td>
<td>Number of deaths</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>K_D_ratio</td>
<td>match</td>
<td>integer</td>
<td>ratio K/D against the killed player, no matter the team</td>
<td>matches/match_id</td>
</tr>
<tr>
<td>M4A1</td>
<td>match</td>
<td>integer</td>
<td>number of sub-rounds named M4A1</td>
<td>matches/match_id</td>
</tr>
</tbody>
</table>

Figure A.1: raw data columns extracted from FACEIT API
<table>
<thead>
<tr>
<th>Feature</th>
<th>Granularity</th>
<th>Description</th>
<th>Comments</th>
<th>endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>player_id</td>
<td>player-frame</td>
<td>integer</td>
<td></td>
<td>playerFrame table</td>
</tr>
<tr>
<td>frame</td>
<td>player-frame</td>
<td>integer</td>
<td></td>
<td>playerFrame table</td>
</tr>
<tr>
<td>match_id</td>
<td>player-frame</td>
<td>integer</td>
<td></td>
<td>playerFrame table</td>
</tr>
<tr>
<td>second</td>
<td>player-frame</td>
<td>float</td>
<td></td>
<td>playerFrame table</td>
</tr>
<tr>
<td>side</td>
<td>player-frame</td>
<td>string</td>
<td>Terrorist or Counter Terrorist</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>health_points</td>
<td>player-frame</td>
<td>integer</td>
<td>Remaining health points</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>armor_points</td>
<td>player-frame</td>
<td>integer</td>
<td>Remaining armor points</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>position</td>
<td>player-frame</td>
<td>string</td>
<td>Geoposition in the CS:GO Map (XY,Z) coordinates</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>grenades</td>
<td>player-frame</td>
<td>bool</td>
<td>If player has grenades (one variable by grenade type)</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>is_alive</td>
<td>player-frame</td>
<td>bool</td>
<td>If player is alive</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>is_bot</td>
<td>player-frame</td>
<td>bool</td>
<td>If player is bot</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>is_planting</td>
<td>player-frame</td>
<td>bool</td>
<td>If is planting a bomb</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>is_defusing</td>
<td>player-frame</td>
<td>bool</td>
<td>If is defusing the bomb</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>cash_spent</td>
<td>player-frame</td>
<td>integer</td>
<td>Cash spent buying equipment</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>ping</td>
<td>player-frame</td>
<td>integer</td>
<td>Connection latency</td>
<td>playerFrame table</td>
</tr>
<tr>
<td>round_number</td>
<td>round</td>
<td>integer</td>
<td></td>
<td>round Table</td>
</tr>
<tr>
<td>match_id</td>
<td>round</td>
<td>integer</td>
<td></td>
<td>round Table</td>
</tr>
<tr>
<td>Terro_score</td>
<td>round</td>
<td>integer</td>
<td>Terrorist score at the end of the round</td>
<td>round Table</td>
</tr>
<tr>
<td>CT_score</td>
<td>round</td>
<td>integer</td>
<td>Counter Terrorist at the end of the round</td>
<td>round Table</td>
</tr>
<tr>
<td>Terro_team_name</td>
<td>round</td>
<td>string</td>
<td>Terrorist team name</td>
<td>round Table</td>
</tr>
<tr>
<td>CT_team_name</td>
<td>round</td>
<td>string</td>
<td>Counter Terrorist name</td>
<td>round Table</td>
</tr>
<tr>
<td>winning_side</td>
<td>round</td>
<td>string</td>
<td>Round winner faction (Terrorist or CT)</td>
<td>round Table</td>
</tr>
<tr>
<td>winning_team</td>
<td>round</td>
<td>string</td>
<td>Winner team name</td>
<td>round Table</td>
</tr>
<tr>
<td>round_end_reason</td>
<td>round</td>
<td>string</td>
<td>Round end reason (Kills, bomb, time...)</td>
<td>round Table</td>
</tr>
</tbody>
</table>

Figure A.2: raw data columns extracted from FACEIT Demofiles
<table>
<thead>
<tr>
<th>Feature</th>
<th>Granularity</th>
<th>Description</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>seconds</td>
<td>integer</td>
<td>millisecond video</td>
</tr>
<tr>
<td>angry</td>
<td>seconds</td>
<td>float</td>
<td>Proportion of the emotion</td>
</tr>
<tr>
<td>happy</td>
<td>seconds</td>
<td>float</td>
<td></td>
</tr>
<tr>
<td>surprise</td>
<td>seconds</td>
<td>float</td>
<td></td>
</tr>
<tr>
<td>neutrality</td>
<td>seconds</td>
<td>float</td>
<td></td>
</tr>
<tr>
<td>fear</td>
<td>seconds</td>
<td>float</td>
<td></td>
</tr>
<tr>
<td>sadness</td>
<td>seconds</td>
<td>float</td>
<td></td>
</tr>
<tr>
<td>disgust</td>
<td>seconds</td>
<td>float</td>
<td></td>
</tr>
</tbody>
</table>

Figure A.3: raw data columns extracted from video files through DeepFace Python library

Figure A.4: No smoothed emotions, Participant 4