POLITECNICO DI TORINO

Master's Degree in COMPUTER ENGINEERING

Master's Degree Thesis

Awe by video games: how do video games elicit awe?

Supervisors

Prof. ELISA MEKLER

Prof. LUIGI DE RUSSIS

Candidate

AURORA ANNA PIA SERGIO

FEB 2023

Abstract

Although overlooked for decades, video games have gained more and more attention as a medium capable not only of entertainment, but also of provoking deep emotions in players. One emotion that seems to be evoked across many different kinds of games is awe. This thesis tries to explore the link between video games and the experience of awe; in particular, the goal is to find which game features are more likely to induce this emotion in players. To do so, a survey is used to collect relevant data; the collected answers are then processed and analyzed by means of decision trees, in order to understand the connection between each feature and awe. Performance scores varied across the different models built from the data, with some of them resulting in a more reliable prediction than others. The former group of models suggest that the features seemed to induce awe are the following: graphics described as disturbing, elaborated plots, epic soundtracks, extraordinary locations (like modern cities or fantasy worlds), and a stimulating amount of challenge. These features can work as directions in the design phase, so that video games professionals can aim at creating awe-inducing products; such choices would lead to beneficial effects in players, like enhanced pro-social behaviours, and might also improve the game's popularity.

Acknowledgements

This master's thesis comes at the end of two important journeys in my life. One is my study course at Politecnico di Torino, which has been full of challenges and satisfaction. The other one is my exchange at Aalto University, which enriched my studies even further and allowed me to try new experiences, uncovering new parts of me.
I first want to thank my supervisor, Professor Elisa Mekler, and my advisor, Professor Luigi De Russis, for their guidance and precious advice, that helped me transform my idea into this thesis. But this work would have not been possible without the precious experiences I had during my university years, that made me grow, change, and learn.
First, I want to thank my oldest and best friend Gaetano, who has been with me since the very first day, always supportive and understanding, always by my side.
For all the fun years in our student residence, I have to thank my friends from Collegio Einaudi, especially Alessio, Alessandro, Antonio, Gianluca, Jacopo, Kevin, and Morena.
Special thanks go to Antonio, for being a trusted and inspiring project partner during our Master's courses, because no matter the difficulties, we always managed to make the best out of our assignments.
Not all of these years took place in Italy, though; I want to thank the coolest Ryhmä ever, for making my exchange so special, and for being by my side all along. These thanks are for Alex, Christian, Dominik, Jana, Jonathan, Luca, Niklas, and Ondřej, without whom my year in Finland would not have been so rich of wonderful moments and experiences. I am grateful I met them, because they made me grow and they encouraged me to overcome my limits, always reassuring me in the process; and even though most of them like pineapple on pizza, they all have a special place in my heart, and always will.

Last, but not least, I want to thank my family, for all the sacrifices they made to raise me and to allow me to study, and for always putting great value on education. I know that they are proud of me and that they love me, and I hope I will keep making them proud even after the end of my studies.

Table of Contents

Chapter 1

Introduction

Since the emotion of awe has been given its first formal definition by Keltner and Haidt [1], research on the topic has worked to better understand it. Awe is a complex emotion, that can have both positive and negative valence, although it is more commonly associated with positive feelings and outcomes. Its fundamental aspects are *perceived vastness* - either physical or conceptual - and a *need for accommodation*, a natural process that occurs when the human brain cannot resolve a stimulus by using the existing mental structures and therefore needs to adapt them or create new ones. Studies have found awe triggers in nature [2], music [3], art [4], people [5], and even virtual reality [6]. It also has shown to have positive effects on well-being, together with increasing pro-social behaviors [7].
In the meantime, video games have been undergoing a process of re-evaluation and became the object of numerous psychological studies, finding benefits related to their use [8]. Starting as a mere form of entertainment, it was in the 1980s that the game industry started claiming the artistic potential of video games. Critics, scholars, and philosophers have been active in the debate ever since; the general sentiment went from disregarding the value of video games to the acknowledgment of video games as a new art form in the mid-2000s [9], although perhaps not comparable with established forms of art. Of course, not all games are, or even aim to be, considerable as art; however, some games are designed to evoke a wide range of emotions in players, effectively becoming closer to art in a conventional sense. Examples of 'emotional' games are *Flow*, *Flower*, and *Journey*, together with the class of 'empathy games', described as games that ask "players to inhabit their character's emotional worlds" [10]. Other studies have investigated the players' experience of video games as art, finding that, among all the emotions that players report feeling when experiencing games as art, there was awe [11].
However, the components of video games that allow them to induce awe have not been investigated yet. The goal of this work is to try to understand how games can be triggers of the emotion in players - in other words, which of their features are

more relevant to induce the experience. Other than filling the gaps in the current understanding of video games as awe elicitors, deeper knowledge in the area can be of interest both for better-comprehending games' impact on players and for guiding game designers in their choices, if their goal is to induce this emotion in players. The study proceeds by collecting data about awe experiences with video games through the distribution of a survey, and then analyzing responses with the aid of machine learning tools, in order to identify which game features are more relevant to the experience.

Chapter 2

Background

This section introduces some concepts essential to the following discussion. First, a brief description of awe is given, so that the reader has a clear idea of what feeling is the object of the study. Then, the goals of this work are presented.

2.1 Awe

Psychologists have spent decades trying to come up with a comprehensive set of fundamental human emotions, or, to be more precise, families of emotions. However, awe has not been recognized as one of them until 2003, when Keltner and Haidt [1] discerned it from the existing families and formalized a prototypical model of awe. According to this model, awe is a complex emotion, characterized by two central elements: *vastness* and a *need for accommodation*. To put this simply, awe occurs when a person perceives something vast, immense, unbounded, either physically or conceptually (e.g. enjoying the view from the top of a mountain, or being presented with the concept of general relativity), that the existing mental schemas cannot fully comprehend - therefore the process of accommodation is needed, which involves adapting or evolving the existing mental structures or building new ones. Unlike most emotions, awe can have an either positive or negative valence, with the latter being less common [12]. Since Keltner and Haidt's definition, awe has been the object of numerous studies, which have found it to be often accompanied by a change in perception of how time flows (often felt as slower or even stopping), a sense of self-diminishment and increased connectedness to other people or nature [13]. Other studies investigated awe elicitors, finding that natural scenarios [2] are one of the most common ones, but also art [4] and music [3], as important mental stimuli, have the potential to trigger the emotion.
On top of the two core elements of vastness and need for accommodation, five additional emotional themes have been identified that produce different awe-related

states; these are *threat*, *beauty*, *ability*, *virtue*, and *supernatural*. The theme is often correlated with the nature of the source of awe.
Different theories have been formulated about the evolutionary function of awe. As with other emotions, it is posited that awe might have addressed a specific survival problem.
Adopting this "functionalist" view on emotion evolution, Kelter and Haidt (2003) [1] suggest that awe might have served the purpose of maintaining social hierarchies. According to this theory, awe was first experienced in relation to powerful leaders, which provoked a sense of fear and respect in subordinates. Therefore awe could have facilitated gatherings of people around charismatic and dominant figures, creating and reinforcing a sense of shared identity. This view implies that awe originated as an emotion in social contexts, and only afterward generalized to include nature-related triggers.
An opposing view, suggested by Chirico and Yaden [12], switches the natural and social contexts: it posits that awe was first triggered by vast natural scenarios, and only afterward the social trigger appeared. The evolutionary advantage of awe from this point of view would be that it allowed finding better shelters: these were characterized both by safety and the possibility to have a clear view of the surroundings, essential to spot potential threats in time. Naturally, elevated locations provide both these elements, and to this day the view of immense natural landscape that can be achieved from such locations is considered one of the main triggers of awe.
There is no general agreement on a single theory for explaining the evolution of awe, and it is out of the scope of this work to debate this further; however, the effects of awe experiences that have been observed in multiple studies are interesting to consider. Awe has the important consequences of diminishing one's sense of self and its concerns, and increasing prosocial behavior - characterized by increased generosity, compassion, ethical decision-making, and decreased entitlement [7]. Other studies have instead focused on nature-induced awe, showing that it improved general well-being and stress-related symptoms even in just one week of exposure [14]. Concerning video games, it has been observed that, if playing commonly produces self-transcendent emotions - including awe [15] - such games could encourage self-improvement and growth [16]. Another interesting result is that awe makes a unique contribution to players' sense of entertainment, possibly even fueling players' entertainment experiences [17], and has the potential to improve game enjoyment and appreciation [18]. All these studies point out the positive effects of feeling awe; if video games designer had a clearer idea about how to make players experience it, they could make it one of their design goals; if so, their games could become valuable means for improving the mentioned aspects in players' lives, and in return the games' popularity might grow, meaning that the economical aspect could potentially benefit from such design choices. However, it is not an extremely

simple task: one of the main problems, when researching awe, is to find a way of inducing that emotion among participants. One effective tool for that is virtual reality [12], but video games cannot rely only on that for creating awe experiences in a larger audience, as it is not widely available yet.

2.2 Goals

This work has the main purpose to find out which features of a game are more likely to make it awe-inducing. If these features could be pointed out, designers will know what aspects are to be prioritized, especially with limited resources, in order to make their game more emotionally powerful. Since the scope is quite vast, as games involve several different areas and each of them could contribute to the emotional experience, the aim is to find general directions, rather than extremely precise elements to include in the game design.
Aspects taken into consideration for this study have been chosen to cover all possible areas of video games that might relate to or contain known triggers of awe: these include, but are not limited to, the graphics, the plot, the soundtrack, the setting, and the use of virtual reality technology. It is expected, based on what the literature suggests, that the graphics will turn out to be relevant because they can easily contain factors that are known to elicit a feeling of awe. For the same reason, natural locations like forests, mountains, the ocean, or space, might also be linked to awe experiences. For example, if the game shows multiple scenes of vast landscapes in high-quality graphics, at least some players will probably experience some nuance of awe.

Chapter 3

Methods

This section discusses how the present study has been conducted. The first step was designing a survey to collect the necessary data. The second step was the analysis of the collected data, supported by machine learning tools.

3.1 The survey

Because data about video games and awe experiences are scarcely available, it was necessary to collect ad hoc data for this study. Among several possible ways of doing so, a survey has been deemed to be the most efficient one, allowing a quicker collection of answers than other methods, such as interviews.
Participants have been recruited mainly via *judgemental sampling* [19], an improved variant of *convenience sampling* in which researchers select, among readily available people, those who are deemed to be most representative for the specific study. This is a non-probabilistic sampling process; although the probabilistic kind should be preferred, it is not always a viable option. In this case, with probabilistic sampling, chances were that interviewed people would have not been able to provide useful data for the research, either because they do not play video games at all, or because they have not experienced any strong emotion connected to video games. Moreover, randomly selected people might not be interested in the topic of the survey, making it harder to reach an adequate number of responses. Therefore, probabilistic sampling would have been hard to apply, which is why a different strategy has been used: participants have been invited to take part in the survey via social network posts in communities centered around video games, as these were more likely to have members interested in the topic and with a personal experience relevant to the study.
For what concerns the sample size, the goal was to reach at least 100 answers, preferably 150-200. Overall, a total of 185 answers were collected. Of these, 13

were given by participants between 15 and 18 years old (7%), 102 by participants between 19 and 26 (48%), 63 between 27 and 35 (34%), and only 20 were 36 or older (11%), with 47 being the highest age reported. The vast majority of participants identified as male (78.9%), while 13% identified as female, 3.8% as non-binary, and the remaining 4.3% preferred not to say. Most respondents (61.1%) reported that they typically spend more than 6 hours per week playing video games, 28.1% play between 2 and 6 hours per week, 5.4% between 1 and 2 hours per week, and the remaining 5.4% plays less than 1 hour per week. Only a minority of participants studies or works (<5% in both cases) in the video games field.
The survey is divided into three sections, described in the following paragraphs; the integral version of the survey can be found in Appendix B.

3.1.1 Introduction to the survey

The first section collects general information about the participant; it also serves the purpose of easing respondents into the survey by asking low-effort questions, such as age or gender. In this section, participants are asked about whether their studies or profession are somehow related to video games, and about how much time per week they spend playing. The reason for these questions is that people with strong background knowledge about video games might appreciate them in a different way from other users; in the same way, someone spending a lot of time playing video games might have a wider range of experiences than people playing only occasionally.

3.1.2 Description of the experience

In the second part, participants are asked to think of a personal awe experience that occurred with video games, and to express their agreement with a battery of statements about the experience. Most statements have been selected from a larger set developed by Yaden et al [13], a six-factor scale intended to measure awe experiences. Not all items from the scale have been used, in order to ease the cognitive demand on participants, an issue that typically occurs in longer surveys and increases the drop-out rate [20]. It has to be noted that, usually, recollecting memories for answering a survey or an interview is not easy, and can lead to *recall bias* [21]. In order to minimize this effect, the instructions at the start of this section encourage respondents to think of a recent experience, write down some notes about it, and take some time to do so. This strategy seems to be effective at helping participants remember emotional experiences, so it is often used in studies about awe [22, 23]. The answers to these questions are later used to distinguish awe experiences from non-awe experiences. As Keltner and Haidt [1] highlighted, there are emotional experiences that are close enough to awe to be easily mistaken for it, but lack one or both the core components of the emotion (*vastness* and *need for accommodation*). An example that well suits the case of video games would be admiration. Quite often, players will notice something impressive that makes them appreciate the game developers' ability, either in the artistic component of video games or in the technical one. However, admiration is a distinct emotion from awe, because it lacks the *vastness* component.

3.1.3 Description of the game

The third section covers the game that elicited the experience, in particular by asking about the following aspects:

- graphics
- plot
- soundtrack
- main character
- genre
- location setting
- pace and difficulty
- use of virtual reality headsets

Participants are asked to write the title of the game, select its genre(s), and then rate some of the game components: these are the graphics, the plot, the soundtrack, and the main character. For each of these, there is also an additional question, asking participants to describe the considered component. For the sake of standardization, each of these questions gives a few elements to select from, but with the additional option to manually type whatever the respondent thinks is suitable to describe the item. After this set of questions, there are some more that cover the other elements of a game: location, pace and difficulty. Respondents are not asked to rate any of these because it is unlikely that someone would have a strong opinion about them. Instead, participants are asked to point out all the locations where the game takes place and to select answers that apply to the game's pace and difficulty. The question about locations is intended to allow identification of the most common ones when it comes to awe experiences, while pace and difficulty might have an impact on the mental state of the player, either facilitating or impeding awe. Finally, the participant is asked whether they were using a virtual reality headset when the experience occurred and if so, they are asked to describe this aspect of the game, similarly to how they have already done with every other aspect in the previous questions.

3.2 Data pre-processing

The first step after collecting responses to the survey is to extract data from them in a suitable format for further elaboration.
Most of the elaboration work is automated and done by a short script created for this purpose (Appendix A contains a link to the GitHub repository), but not completely: since it is not possible to auto-correct all typos, this task was done manually. It did not take much time, as the correction was only relevant for the game titles and a few other fields that were freely completed by participants (e.g. indicating locations that were not listed among the possible answers, or describing any of the video game components). However, it was a useful step to take before proceeding, since responses containing the same keywords had to be grouped together. The spelling correction has been done on a copy of the original data, in order to keep participants' answers intact for possible future use.
In a later stage of the analysis, other changes to text answers were made necessary for better visualization and understanding of results: many answers at this point were summarized in 2-3 words. The edited answers are mostly about the main character, the pace and the difficulty of the game, and the editing involves both the survey's pre-made answers and answers written by participants.
The answers to the survey are saved in a .tsv file; these are very similar to .csv files, except for the separator character (a tab instead of a comma). As the data is read by the program, answers are saved in a `pandas.DataFrame` object - defined as 'a 2-dimensional labeled data structure with columns of potentially different types' in the documentation [24]. Each row of the dataframe corresponds to one answer, each column contains the answers to one question.
The next paragraphs discuss how the program further elaborates participants' answers, preparing them for the machine learning model.

3.2.1 Identifying and removing low-effort responses

One of the main problems with surveys is that some participants might fill them out until the end, but without investing enough mental energy in the task. The result is that the data collected from these participants are low in quality and could affect analysis negatively. Although it is inevitable to get such respondents in a survey study, there are ways to spot those so-called "lazy" participants, and their answers can be excluded from the study.
In the survey used for this study, there were two main indicators for recognizing low-effort responses: first, in section two of the survey, those responses were made identifiable by "straight-lining", a behavior that consists in choosing always the same answer across multiple questions; second, in section three of the survey, the answers to three questions called for one specific answer in another three questions,

therefore participants who did not give the expected answers in the second group of questions were either answering “randomly" to the survey or did not pay enough attention for their answers to be considered valid. The questions were numbers 26, 27, 30, 31, 34, and 35 in the survey (see Appendix B).
The following code works on the `dataframe df`, collecting all the answers to the survey, one per row. The first method, `remove_lazy()`, calls two more methods per row: with the first one, `lazy()`, straight-lining is detected; with the second one, `check_pairs()`, incompatible answers are detected.

```
import numpy as np
import pandas as pd

def remove_lazy(df):
    indexes = []  # track rows to remove
    num_rows = df.shape[0]  # rows in df
    for i in range(0, num_rows): # for each row
        if lazy(df.iloc[i]): # if straight-lining
            indexes.append(i)   # add row index
            continue
        if not check_pairs(df.iloc[i]): # if random answers
            indexes.append(i)
    df = df.drop(indexes, axis=0)   # remove tracked indexes
    return df


def lazy(row):
    answer_streak = 1  # count how many consecutive answers are the same
    for i in range(6, 20):  # 6 to 20 = answers in section 2
        if row[i] == row[i - 1]:
            answer_streak = answer_streak + 1
            if answer_streak >= 8:  # streak 8/15, the respondent is considered lazy
                return True
    return False


def check_pairs(row):
    # if the rating is empty, the description should be 'not applicable'
    # if the rating is not empty, the description should not be 'not applicable'
    if row['Story_rating'] == np.NaN and row['Story_descr'] != 'not applicable':
        return False
```

```
    if row['Story_rating'] != np.NaN and row['Story_descr'] == 'not applicable':
        return False
    if row['Main_character_rating'] == np.NaN and row['Main_char_descr'] != 'Not applicable':
        return False
    if row['Main_character_rating'] != np.NaN and row['Main_char_descr'] == 'Not applicable':
        return False
    if row['VR'] == 'No' and row['VR_descr'] != 'not applicable':
        return False
    if row['VR'] == 'Yes' and row['VR_descr'] == 'not applicable':
        return False

    return True
```

With these criteria, out of the 185 collected answers, 140 were kept for further analysis.

3.2.2 Generating label column for supervised machine learning

Machine learning algorithms can be first split into two categories: supervised and unsupervised [25]. The main difference between the two is that supervised algorithms work on 'labeled' data. This means that the quantity that the machine learning model is trying to predict is already known, at least on the training data, and can be used to measure the model's accuracy. Unsupervised algorithms, instead, explore the data to find hidden patterns, without the need for human intervention. Supervised methods are further categorized in regression and classification. The difference between regression and classification problems is in what the model predicts: for regression problems, it is usually a number in a continuous range (e.g. the maximum temperature in a day); for classification problems, the model predicts one out of a finite set of categories, or classes (e.g. a picture represents a cat or a dog). Because unsupervised algorithms are designed to work on large amounts of data, are very complex, and require powerful tools, while supervised algorithms are well suited for predicting outcomes for new data, and tend to be simpler, the latter category is chosen to work on this problem. The problem could be formulated like this:
"Given a data point, containing information about a player and a game, will that game elicit awe in that player?"
The possible answers to this question are not in a continuous range: in fact, they are limited to "Yes" and "No". A data point can either represent an awe experience

or not. Therefore, it is clear that the problem falls into the classification category. At this point, the data needs to be labeled. To do so, the answers in section two of the survey are used to distinguish between participants who actually reported an awe experience and those who might have experienced something different. The questions in this part of the survey were chosen to be highly correlated with awe experiences. Therefore, if respondents indicate strong disagreement with a majority of those statements, meaning that they did not experience the described feeling or mental state at all, then their experience can be considered a 'non-awe' experience. Two items in particular, indicating that the participant felt the mentioned need for accommodation, were checked for each answer. If the respondent reports a strong disagreement with any of those two statements, it means that one of the fundamental components of awe was not experienced at all. For this reason, this kind of answer is also marked as a 'non-awe' experience. The code used for creating the label to add to the `dataframe df` is the following:

```
def create_label(df):
    labels = []  # list that will become the new column
    num_rows = df.shape[0]

    for i in range(0, num_rows):
        labels.append(is_awe(df.iloc[i]))  # append 1 if this row corresponds to an awe experience, 0 otherwise

    df['Felt_awe'] = labels  # add new column to the dataframe

    drop_awe = ['Time_change', 'Slowed_down', 'Smaller_self', 'Oneness_things', 'Humbling', 'Connected_humanity', 'Something_greater', 'Chills_goosebumps', 'All_at_once_struggle', 'Mentally_challenged', 'Positive_impact', 'Fear_discomfort', 'Peace_of_mind', 'Admiration_game_dev', 'Better_person']
    df = drop_irrelevant_columns(df, drop_awe)  # the columns about awe are not needed anymore
    return df


def drop_irrelevant_columns(df, col):
    # remove columns that are not useful from the dataframe
    df = df.drop(col, axis=1)
    return df


def is_awe(row):
```

```
    # check if a row of the dataset corresponds to an awe experience (1) or not (0)
    if row['Mentally_challenged'] == 1 or row['All_at_once_struggle'] == 1:  # main component of awe is not experienced
        return 0

    count = 0  # how many of the important items were rated 1 ('Strongly disagree')
    for i in range(0, 12):  # the first 12 columns *of the second section* measure awe
        if row[i + 5] == 1:  # i + 5 to skip the answers of the first section
            count = count + 1
    # if we got to the for loop, it means neither of mentally_challenged nor struggle were ones, so we only have to check the remaining 10 answers. Therefore, if 6 or more of these aspects were rated as ones, the experience was probably not awe
    if count >= 6:
        return 0
    return 1  # passed both checks
```

3.2.3 Adjusting text answers and interpreting NaN values

The method discussed here is used to make minor changes to text answers, in order to make them easier to handle in the following steps. In addition to that, some answers are replaced with values that are more suited for the analysis. The applied changes are:

- Removing the examples given for genres in parentheses
- Removing the nonessential elements between commas, like “at least partially" and “to a certain extent" in the answers about the main character. These phrases were in the answers for legibility and ease of understanding, but are not fundamental to the sentences
- Switching some commas for other signs of punctuation. In the following part of the processing, commas will be used to identify and separate different items in a list; this step removes those commas that are not splitting items, but are contained within one item (e.g. “woods, forests" is turned into “woods/forests")
- Replacing Yes/No answers with True/False values. This applies to the questions about the participant studying or working in a video game related field, and about the use of a virtual reality headset when the reported experience occurred

- Replacing missing values in the “story rating" and in the “main character rating" columns. These correspond to the two questions asking the respondent to give their rating about the story and the main character in the game that elicited their experience but left the option to skip them, respectively if the game did not have a proper story or if it did not involve controlling a character. A blank answer would mean that the feature has little relevance to the outcome (e.g. even without a main character, the game managed to provoke awe), as opposed to a high rating suggesting that extremely good features of a game might elicit stronger feelings in the players. For these reasons, missing values are replaced with zeros.

```python
def replace(df):
    # first, remove the parentheses because they were just supposed to help respondents
    df.Genre = df.Genre.str.replace(' \([^)]*\)', '', regex=True)

    # second, convert yes and no into true and false (so we already have boolean types in some answers)
    df = df.replace({'Study_videogames': 'Yes', 'Work_videogames': 'Yes', 'VR': 'Yes'}, True)
    df = df.replace({'Study_videogames': 'No', 'Work_videogames': 'No', 'VR': 'No'}, False)

    # now removing the "between commas" sections, which were there for legibility
    df.Main_char_descr = df.Main_char_descr.str.replace(', at least partially,', ' ', regex=True)
    df.Main_char_descr = df.Main_char_descr.str.replace(', to a certain extent', ' ', regex=True)
    df.Story_descr = df.Story_descr.str.replace('weak, contains plot holes', 'weak and containing plot holes')
    df.Soundtrack_descr = df.Soundtrack_descr.str.replace('on spot, perfect for the game', 'on spot; perfect for the game')
    df.Soundtrack_descr = df.Soundtrack_descr.str.replace('irrelevant, left me indifferent', 'irrelevant; left me indifferent')
    df.Locations = df.Locations.str.replace('woods, forests', 'woods/forests')
    df.Locations = df.Locations.str.replace('sea, ocean', 'sea/ocean')
    df.Locations = df.Locations.str.replace('space, spacecraft', 'space/spacecraft')
    df.Pace_and_difficulty = df.Pace_and_difficulty.str.replace('the game was challenging, sometimes too much', 'challenging')

```

```
    # finally, replace NaNs:
    df = df.replace({'Story_rating': np.nan, 'Main_character_rating':
    ↪  np.nan}, 0)
    df = df.astype({'Story_rating': np.int64, 'Main_character_rating':
    ↪  np.int64})

    return df
```

3.2.4 Getting separate dataframes for ratings and descriptions

From the answers to this survey, the number of possible features to consider is already quite high. However, most of them need further splitting: for example, one of the questions asked the participant to list all the locations of a game; these answers need to be split into multiple features so that every single location corresponds to one feature (with its own column in the dataframe). The result would be a number of features that is way too high for the available data. For this reason, the data is split into several smaller dataframes, grouping together features that are related to each other; the models, discussed in the following sections, will train on all of them, and the most important features from each dataframe will be extracted.
The first of these smaller dataframes is built by the method `get_ratings(source, cols)`, differently from all the others. The method selects some columns from the main dataframe: these contain the answers about whether work and studies are related to video games, together with all the answers about ratings, and whether VR headsets were used. At this point, a consideration about the answers needs to be made: different respondents have had awe experiences with the same games, therefore some titles are mentioned in multiple rows of the dataframe. For this reason, grouping answers by game title was considered; however, because different players can have different opinions (and evaluate the same aspects of games differently from each other), it was resolved to leave all answers separated, even if they referred to the same game. Another reason for this decision is that, while ratings could easily be merged by averaging across answers about the same title, the same operation cannot be performed for all answers (e.g. how can the playing frequency be merged for different respondents? Would it even make sense?) - which might have created inconsistencies in the dataframe. After selecting the columns from the dataframe, the method also performs one-hot encoding of the playing frequency, by calling `pd.get_dummies()`. One-hot encoding is a technique for converting categorical data, which consists in creating a new column for each unique value in the original category. Then, each new column contains a 1 if the original row contained that value, and 0 otherwise. As an example to further clarify how one-hot encoding works, consider the playing frequency: there were only four

possible answers - less than 1 hour per week, 1-2 hours per week, up to 6 hours per week, and more than 6 hours per week. One-hot encoding, therefore, creates four columns: the first one will contain a one only for players that reported playing less than 1 hour per week, and zero in all other rows; the second one will contain a one for players that answered “1-2 hours per week", and zero everywhere else; and so on.

The other answers were collected into smaller dataframes by the method `get_descrition_df(df, column)`. This works on all the answers to the questions along the lines of ‘how would you describe x? Select all that apply’, together with the answers about the locations, the pace, and the difficulty of the game. These answers contain a list of items separated by commas. Therefore, the method first splits those lines into a list of items and then makes use of an object, provided by the `sklearn` library, called `MultiLabelBinarizer`. This is used to perform one-hot encoding on the items of the list, so that, to each item, corresponds a new column. The method offered by `pandas`, `get_dummies()`, previously used for one-hot encoding, would not have worked in the same way on lists, so another solution had to be used.

```python
from sklearn.preprocessing import MultiLabelBinarizer

if __name__ == '__main__':
    # [previous code]

    # split dataset in ratings + demography vs other data
    ratings_names = ['Study_videogames', 'Work_videogames', 'Graphics_rating', 'Story_rating', 'Soundtrack_rating', 'Main_character_rating', 'VR']
    ratings = get_ratings(awe_data, ratings_names)
    awe_data = drop_irrelevant_columns(awe_data, ratings_names)

    genre = get_description_df(awe_data, 'Genre')
    graphics_description = get_description_df(awe_data, 'Graphics_descr')
    story_description = get_description_df(awe_data, 'Story_descr')
    soundtrack_description = get_description_df(awe_data, 'Soundtrack_descr')
    main_char_description = get_description_df(awe_data, 'Main_char_descr')
    locations = get_description_df(awe_data, 'Locations')
    pace_diff = get_description_df(awe_data, 'Pace_and_difficulty')
    vr_descr = get_description_df(awe_data, 'VR_descr')

    # [more code]
```

```python


def get_ratings(source, cols):
    res = pd.DataFrame(source[cols])

    objects = ['Play_frequency']
    for obj in objects:
        awe_data[obj] = awe_data[obj].astype('category')

    # add one-hot encoded play frequency
    res = pd.concat([res, pd.get_dummies(source.iloc[:, 2], prefix='plays')], axis=1)
    return res


def get_description_df(df, column):
    # order items alphabetically and make lists out of them
    df[column] = [','.join(sorted(i.split(', '))) for i in df[column]]
    df[column] = df[column].str.split(',')  # to make a list again

    # one hot encoding
    mlb = MultiLabelBinarizer()
    res = pd.DataFrame()
    tmp = pd.DataFrame(mlb.fit_transform(df[column]), columns=mlb.classes_, index=df.index)
    res = pd.concat([res, tmp], axis=1)

    return res
```

3.3 Decision Trees for Feature Selection

Of the possible machine learning algorithms that can be used for classification problems, the decision tree one has been selected for this study.
Decision trees predict the value of a target variable by learning simple decision rules inferred from the data features [26]. The algorithm builds a tree-like structure, with nodes - from which edges leading to more nodes originate - a root - which is the starting node of the tree - and leaves - which are nodes from which no other nodes are created. Leaves usually contain the final decision (i.e. the label), while middle nodes contain decision rules. The algorithm classifies each data point by going through the decision tree: at each node, the data point features are used according to the rule defined for that node, in order to select the next node among the children (i.e. the nodes reachable from the current one via the edges originating

from it). The selected child then becomes the current node, and the process is repeated until the deepest level of the tree is reached. At the second deepest level of the tree, when children nodes are the leaves of the tree, the rules allow selecting the label.
It has to be noted that the term “decision tree algorithm" can refer to one of several algorithms: some examples are ID3 (Iterative Dichotomiser 3), C4.5 (successor of ID3), and CART (Classification and Regression Trees). The implementation of the algorithm used in this study is the one included in the `scikit-learn` library [27], which uses, according to the documentation [26], “an optimized version of the CART algorithm". The trees built by this algorithm are binary trees (each node has only two children, with an obvious exception for the leaves).
The main advantage of using decision trees for this problem is that they are easy to understand and interpret [26]. Moreover, they are well suited for the goal of this study, as they implicitly perform feature selection [28]: this way, it is possible to find out which, among the answers given by participants, are the most predictive of awe experiences. Another advantage is the use of a white box model [29] - as opposed to algorithms that make use of black box models, like artificial neural networks, which yield results that are more difficult to interpret. Other strengths of decision trees, although less relevant in this context, are the reduced need for data preparation, the logarithmic complexity in the number of data points (which makes them well-suited for large datasets), the ability to handle multi-output problems, and the ability to handle both numerical and categorical data [26].
However, as with every other model, decision trees also have some limitations that should be taken into account.
The first, very important problem, is overfitting [26]. Overfitting is a situation that occurs with various machine learning algorithms. It means that the model fits the training data too well - in fact, the model fits the training data so perfectly that it does not generalize well anymore, so it will not be able to make good predictions on different data. Figure 3.1 gives a good representation of the phenomenon.

Luckily, overfitting can be spotted by looking at some simple metrics of the performance of the model, like training and validation accuracy. These represent the fraction of correct predictions over the total number of predictions, respectively on the training portion of the dataset (typically 70-90% of data points are used for training the model), and on the validation set (the remaining data points that have not been used in the training). When the training accuracy is much better than the validation accuracy, it means that the model is probably overfitting. Overfitting can be countered by adjusting some parameters of the model; in the case of decision trees, these parameters are the maximum depth of the tree, or the minimum number of samples required at a leaf node [26].
Another important problem with decision trees is that they might be biased in favor of the dominant class in the dataset. Therefore, if the dataset is unbalanced,

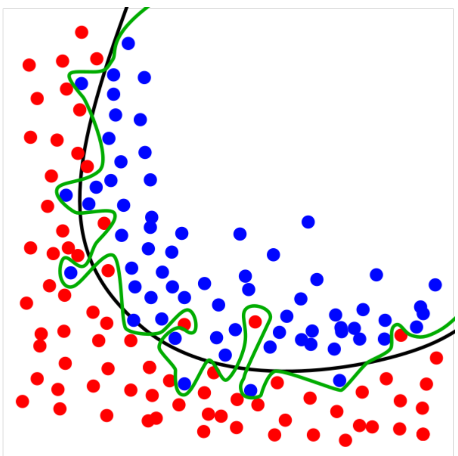

Figure 3.1: Visual representation of ideal fitting (black curve) vs overfitting (green curve), from Ignacio Icke's own work [30]

with samples of one class way more frequent than others, it is better to balance it prior to fitting the tree. The `DecisionTreeClassifier` implemented in the `scikit-learn` library accepts class weights as an optional parameter, useful in these situations.
One other problem of decision trees is the possible instability, as small variations in the initial dataset might produce completely different trees. Finally, the problem of learning an optimal decision tree is known to be NP-complete [26]. Therefore, practical decision-tree learning algorithms tend to adopt different strategies, such as the greedy algorithm, consisting in making locally optimal decisions at each node. These algorithms are necessary to make the problem computationally treatable, but they cannot guarantee to return the globally optimal decision tree.
Since the advantages discussed outweigh the disadvantages of decision trees in this specific context, they still seem to constitute a good tool for the analysis.

3.3.1 Implemetation

At this stage, the actual analysis of the collected answers can start. It can be broken down into two main steps: in the first step, several slightly different models are created and trained on each of the smaller dataframes, in order to tune the model's parameters; in the second step, the best model for each dataframe is used to retrieve the most important features from each of them.
The task of creating different models is done by the method `decision_trees(...)`, called several times in order to process all the smaller dataframes previously created. It receives as parameters the dataframe containing the features, the `pandas series` [31] containing the label, and an optional criterion. The third parameter was used in a first stage of the analysis to experiment with the entropy criterion, too. By default, `sci-kit learn` decision trees use the Gini impurity criterion. The difference between the two is in the mathematical formula of impurity to be minimized, to select the (locally) optimal split of a node. Since the two criteria did not yield significantly different results, and the entropy criterion is more computationally expensive, the third parameter was later not used anymore; therefore, all the trees were created by using Gini impurity.
Before starting to train the models, the method performs an 8-fold split, rather than a simple training set vs validation set split. The choice of using k-fold splitting is due to the reduced dimension of the dataset - 140 data points are few enough to risk incurring in a lucky (or unlucky) split situation. The k-fold cross-validation technique [32] splits a dataset into k smaller sets; then, a single model is trained using k-1 sets, and the remaining one is used for validation; this step is repeated k times, always switching to a different subset for validation and using the other k-1 for training, until all sets have been used as validation set once. The model's accuracy can then be averaged across all k iterations, in order to get a better sense

of how well the model is actually performing. If instead, a simple split between training and validation set was to be used, since the dataset does not have many data points, it could have happened that the dataset was split in such a way that the training accuracy and the validation accuracy appeared to be much higher than what they actually were.
The execution flow then enters a for loop, in which the following steps are performed: first, a `DecisionTreeClassifier` object is created, with a varying maximum depth set by the iteration number, and always with the optional parameter `class_weight` set to balanced, to counter the fact that the class of awe experiences (labeled with a 1) is more common in the dataset, with 108 data points out of 140 falling in this category; then, the model is trained using the k-fold cross validation technique, so the training and validation accuracy values are stored for each iteration of the innermost for loop; before the next iteration, the average training and validation accuracy at the given depth is calculated and printed on the screen. Upon seeing the results, the ideal depth for each of the smaller dataframes is selected; this is usually the depth such that the validation accuracy reaches its maximum value, before it starts decreasing again, and such that the training accuracy and the validation accuracy are as close as possible - to prevent overfitting. The implementation of this method is the following:

```python
import sklearn.tree
from sklearn.metrics import accuracy_score
from sklearn.model_selection import KFold


def decision_trees(df, y, criterion='gini'):

    n = df.shape[1]     # number of features
    tr_accuracies = {}
    val_accuracies = {}
    X = df.iloc[:, :n].to_numpy()    # features
    y = y.to_numpy()    # label
    kf = KFold(8)

    for i in range(1, 10):      # explore max depth
        clf = sklearn.tree.DecisionTreeClassifier(criterion=criterion, max_depth=i, random_state=12, class_weight='balanced')
        tr_accuracies[i] = []
        val_accuracies[i] = []
        for train_index, test_index in kf.split(X, y):
            X_train, y_train, X_test, y_test = X[train_index], y[train_index], X[test_index], y[test_index]
            clf.fit(X_train, y_train)
```

```python
            y_pred = clf.predict(X_train)
            y_test_pred = clf.predict(X_test)

            train_acc = accuracy_score(y_train, y_pred)
            val_acc = accuracy_score(y_test, y_test_pred)
            tr_accuracies[i].append(train_acc)
            val_accuracies[i].append(val_acc)

        avg_tr_acc = np.average(tr_accuracies[i])
        avg_val_acc = np.average(val_accuracies[i])
        print('Average training accuracy at depth ', i, ': ', avg_tr_acc)
        print('Average validation accuracy at depth ', i, ': ', avg_val_acc)
    return
```

The results of this step are summarized in Figure 3.2.

A slight variation of this method is used only for the first dataframe, the one

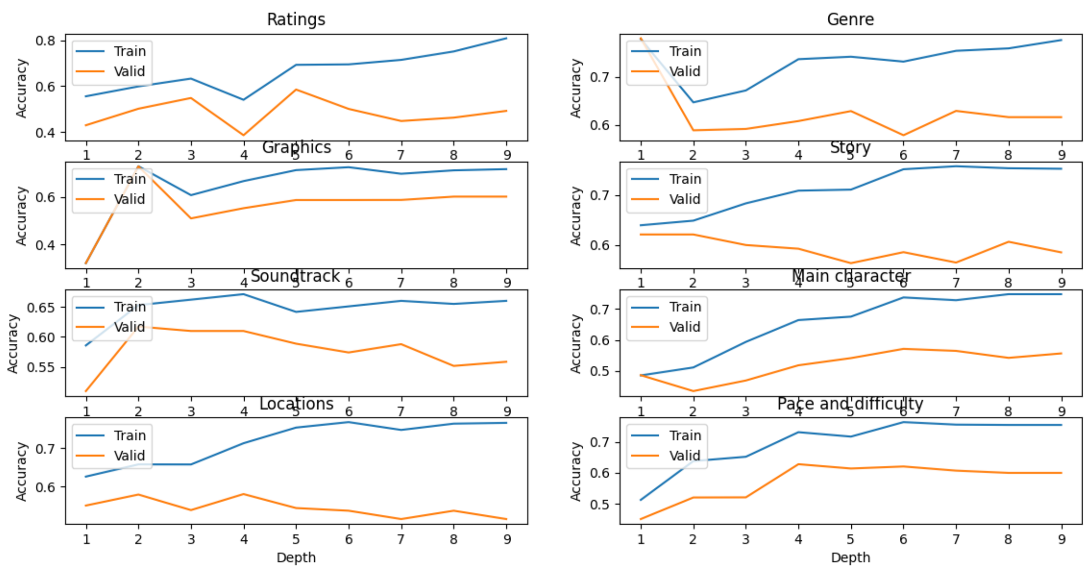

Figure 3.2: Plots of training and validation accuracy for each dataframe

containing the ratings. In this case, the performance of the trees was very poor, so a different strategy has been attempted: instead of trying different depths, different numbers (between 1 and 14) were used as values for the `min_samples_leaf` parameter. This parameter is used to set the minimum number of samples that have to be in any leaf node. The results of this different approach are depicted in Figure 3.3. Although there were a few more trees performing rather badly at any depth, as shown in Figure 3.2, different approaches did not seem to improve the performance significantly, therefore the main one has been used.

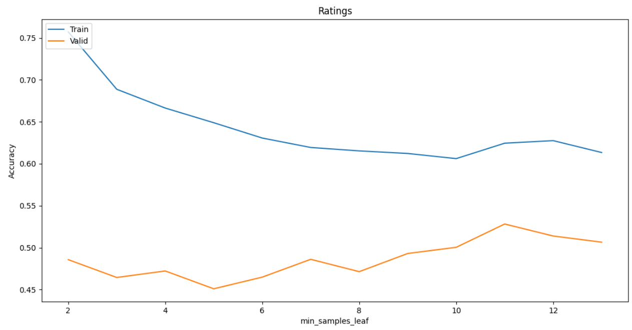

Figure 3.3: Plots of training and validation accuracy for the ratings dataframe

The second step is to use the information learned in the previous one to identify which are the most informative features in each dataframe. To do so, the method `final_decision_tree(...)` receives a dataframe containing features, a pandas series containing the label, an integer to set as the maximum depth of the decision tree, and another to set as the minimum number of samples per leaf, defaulting to 1. The classifier is created and fit as before, but the method performs other significant operations after the classifier has been trained. These operations involve the use of a class, specifically one acting as a meta-transformer for selecting features, `sklearn.feature_selection.SelectFromModel` [33]. The class works alongside any estimator - in this case, the decision tree classifier - that assigns importance to each feature. If the feature importance value assigned by the estimator does not reach a given threshold, the feature is considered unimportant and removed. In this case, the threshold is set to be 1.5 times the mean of all feature importance values. Additionally, it is possible to set a `max_features` parameter to limit the number of features to be selected; in this case, it is set to be equal to the depth of the tree. This way, the dataframes that are used by deeper trees will contribute with more features (as a deeper tree will use more features to split more nodes); on the other hand, when the depth of the tree is low, the number of features actually used is also low, so there is no need for a higher limit. The method then selects the most important features and their names, calculates the training accuracy and validation accuracy of the tree, prints them on screen, and returns a dataframe containing those features. It also plots the decision tree; this is useful to see how the features are used to make decisions, and therefore how their values influence

the prediction.
The method is implemented as follows:

```
from sklearn.model_selection import train_test_split
from sklearn.feature_selection import SelectFromModel
from sklearn.tree import plot_tree
import matplotlib.pyplot as plt

def final_decision_tree(df, y, depth, msl=1):
    # creates the tree at the best depth, returns most important
    ↪ features
    n = df.shape[1]
    X = df.iloc[:, :n].to_numpy()
    y = y.to_numpy()
    X_train, X_test, y_train, y_test = train_test_split(X, y,
    ↪ test_size=0.2, random_state=12)

    clf = DecisionTreeClassifier(max_depth=depth, min_samples_leaf=msl,
    ↪ random_state=12, class_weight='balanced')
    clf.fit(X_train, y_train)

    model = SelectFromModel(clf, prefit=True, max_features=depth,
    ↪ threshold="1.5*mean")
    X_new = model.transform(X)      # reduce X to selected features
    support = model.get_support(indices=True)       # get indices of
    ↪ selected features
    features_names = []
    for i in support:
        features_names.append(df.columns[i])        # get names of
        ↪ selected features for columns

    important_features = pd.DataFrame(X_new, columns=features_names)
    ↪ # df with selected features

    plt.figure(figsize=[12.8, 7.2])
    plot_tree(clf, feature_names=df.columns, class_names=["Non awe",
    ↪ "Awe"], label='none', impurity=False,
              filled=True, proportion=True, rounded=True)
    plt.show()
    print(important_features.info())
    print("Training accuracy: ", accuracy_score(y_train,
    ↪ clf.predict(X_train)))
    print("Validation accuracy: ", accuracy_score(y_test,
    ↪ clf.predict(X_test)))
```

```python
    return important_features
```

Chapter 4

Results

The previous analysis led to the creation and training of several decision trees, each of which uses different features to make a prediction of awe experiences. The way these trees make decisions is what appears interesting at this stage, and this section will proceed to analyse all of them individually. Before approaching the presentation of results, a note about the representation of the decision trees: all the pictures have been plotted by the `plot_tree()`[34] method included in the scikit-learn library. The color of each node is on a spectrum from red to blue, passing through white: a stronger red shade means that the data points that reach the node are more likely to be classified as 'non-awe' experiences, and vice versa a blue shade indicates that the most likely label is 'awe'; nodes that have a white or almost white background color are not very confident in their prediction. Other information included in the nodes is the percentage of data points that reach that node, and the distribution of those data points among the two classes (already balanced).
Before starting, it is necessary to specify that, at this stage, the data related to virtual reality experiences were discarded; the reason for this choice is that there were so few answers from people that were actually using VR headsets, that they were not very informative, and their use would not have further enriched the discussion.

4.1 Ratings

The dataframe containing the ratings given by players to the main aspects of video games, together with some information about the players' habits, was the first one to be studied. The decision tree that was built on these data is shown in Figure 4.1.

What can be seen from the tree is that the first split is done on the rating given

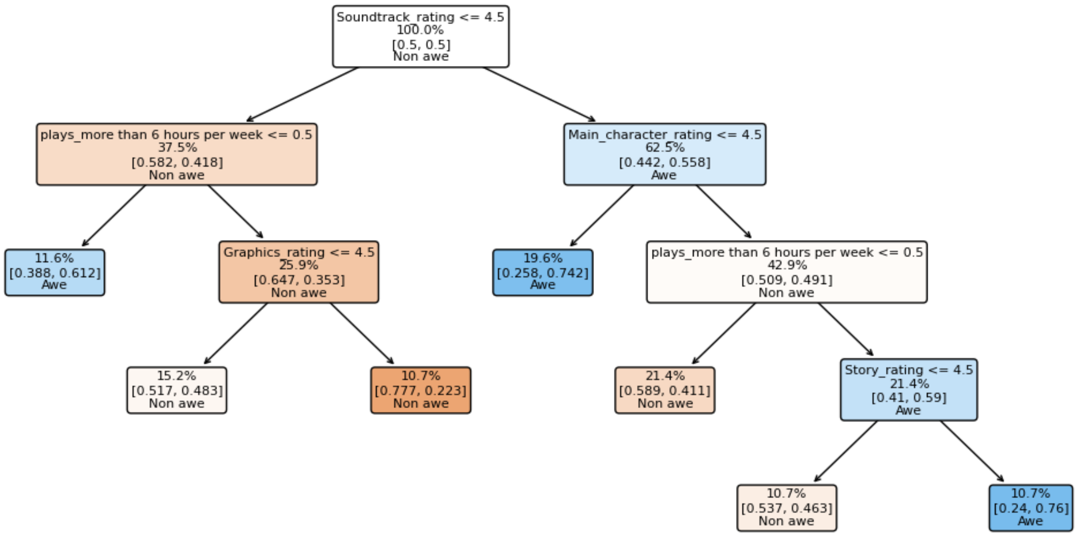

Figure 4.1: Decision tree for the 'ratings' dataframe

to the soundtrack: a rating from 1 to 4 results in fewer awe experiences than a 5. The right child node evaluates the main character rating; curiously, it seems like a rating between 0 (no playable main character) to 4 is linked to a few awe experiences. On both sides of the tree, the playing frequency over 6 hours a week is considered; on the left side, it is negatively correlated with awe; on the right side, instead, it appears to lead to more awe experiences. On the left side, one last split is performed on the graphics rating, where a 5 leads to the data point being labeled as 'non-awe'. The last split, on the far right side of the tree, is done on the story rating, with a 5 leading to an 'awe' label. However, the performance of this single tree is low, as the training accuracy is around 55.4% and validation accuracy is exactly 50%.

4.2 Genre

The next dataframe contains the genre(s) of each game, one-hot encoded. The corresponding decision tree is depicted in Figure 4.2.

This tree did not seem to increase its performance at greater depth. Instead, it

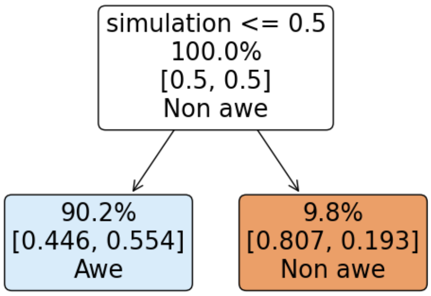

Figure 4.2: Decision tree for the ‘genre’ dataframe

can perform a rather good split based on one genre alone, the simulation one. It appears that simulation games do not produce many awe experiences. This tree, although simple, seems to perform rather well, with training accuracy of 78.6% and validation accuracy of 75.0%.

4.3 Graphics description

The next dataframe to consider is the one containing all the items describing the graphics of games, according to the players. The best performing tree is shown in Figure 4.3.

What can be inferred from this tree is that, when the graphics are considered

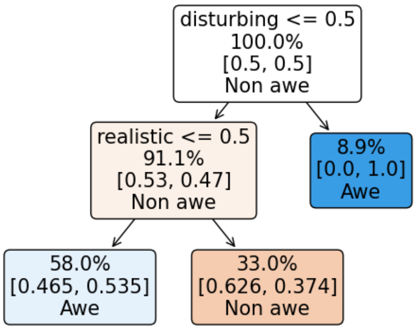

Figure 4.3: Decision tree for the ‘graphics description’ dataframe

‘disturbing’, an awe experience is very likely to be triggered; less than 10% of the data points fall into this case. When the disturbing element is missing, the distinction between the classes is not very clear. It would seem like realistic graphics might not lead to many awe experiences, but the algorithm does not seem strongly confident in the prediction on either side of this node. However, the performance of this tree is also far from ideal, as the training accuracy is around 66.1% and the validation accuracy is roughly 53.6%, and even alternative strategies were not effectively improving it.

4.4 Story description

Some interesting insights are given by the decision tree that was trained on the dataframe containing descriptions of each game’s plot.

Figure 4.4 shows that the first split is made on the attribute ‘rich in subplots’; it

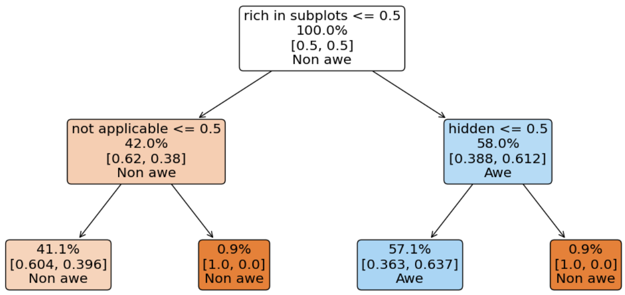

Figure 4.4: Decision tree for the ‘story description’ dataframe

seems that most of the time, this element is directly correlated with awe experiences. Unfortunately, each of the splits at the following levels of the trees only allows precise classification of one single item - apparently, the two participants who answered 'not applicable' and 'hidden' did not report an awe experience. Although these two splits are not informative, again the use of alternative approaches, like increasing the depth or using the `min_samples_leaf` parameter, did not yield significantly better performances. The result does not seem to be a very clear distinction between the two classes: however, the accuracy of the prediction of this tree is 63.3% for the training set and 67.9% for the validation set.

4.5 Soundtrack description

Figure 4.5 shows how the decision tree trained on the soundtrack dataframe uses those features to classify data points.

The first split is performed on the pleasantness of the soundtrack, although it

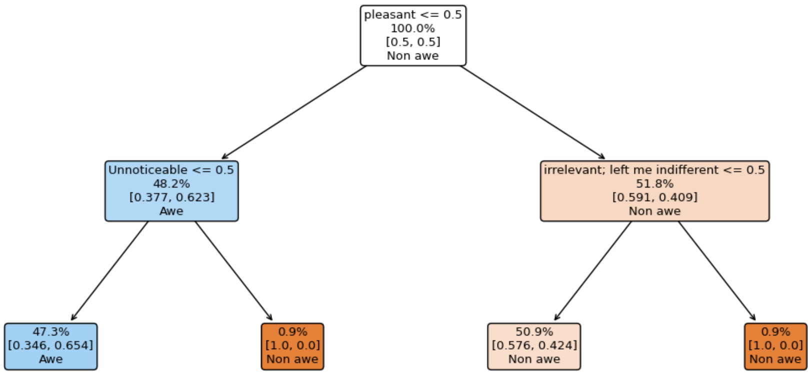

Figure 4.5: Decision tree for the 'soundtrack description' dataframe

might seem a bit counter-intuitive. The second level of the tree performs splits that only allow precise classification of two data points of 'non-awe' experiences. On one side, soundtracks that have not been described as pleasant (which does not necessarily mean they were unpleasant, just that participants did not feel like the adjective suited them), that are also *not* unnoticeable, seem to lead to a decent number of awe experiences. On the other side, music that does not capture the player's attention, and does not contribute to the experience of the game, seems to be associated with non-awe experiences. Both training accuracy and validation accuracy for this model are about 57.1%.

4.6 Main character description

The dataframe discussed next contains information about the main character. The

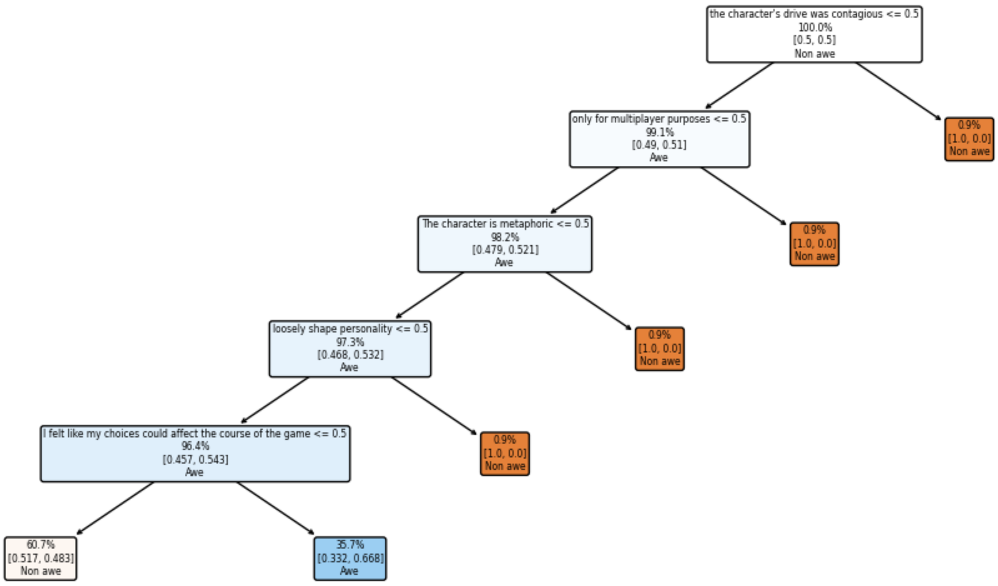

Figure 4.6: Decision tree for the 'main character description' dataframe

tree in Figure 4.6 has a peculiar structure: the first four levels are basically used to classify single non-awe experiences. The fifth node is the only one performing a split between two groups, over whether the player reported feeling like their choices could shape the course of the game or not. The split does not yield a very confident prediction, but performs a bit better on the right side - meaning that, most of the time, when the players feel agency in the game, they also experience awe. However, on the left side the two classes are almost balanced, meaning that when players do not perceive their own agency, no conclusion can be drawn.
The training accuracy of the tree is 49.1%, and the validation accuracy is 53.6%; also in this case, different approaches have been tried, with no significant improvement.

4.7 Locations

Next, the decision tree trained on the locations dataframe can be observed.
By looking at Figure 4.7, it can be seen that the games set in space generally

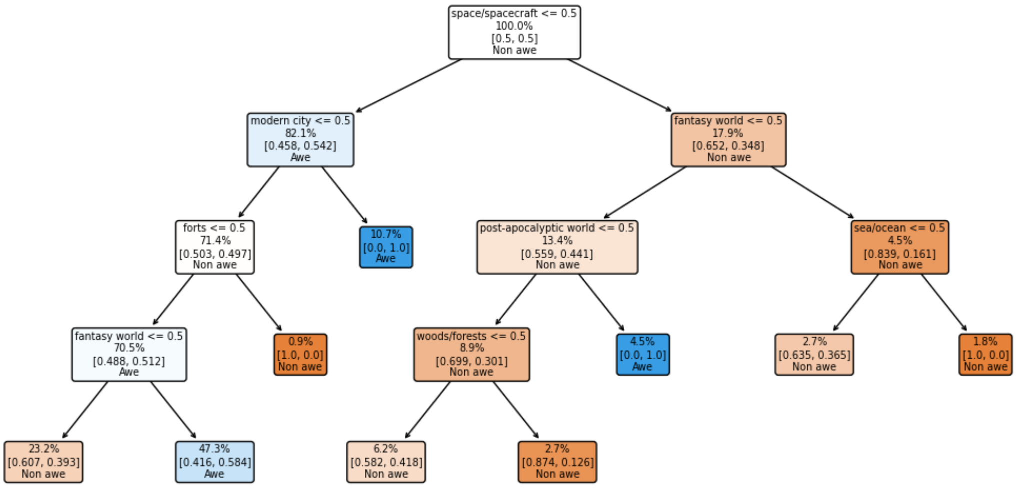

Figure 4.7: Decision tree for the ‘locations’ dataframe

lead to fewer awe experiences; an exception seems to be constituted by games set in space and in a post-apocalyptic world (4.5% of the data points). On the other hand, if space is not one of the settings, awe tends to occur if the game is set either in modern cities or fantasy worlds. Here, the training accuracy is about 68.8% and the validation accuracy is 67.8%.

4.8 Pace and difficulty

The last subset of features considered includes information about the pace and the difficulty of each game. The corresponding decision tree is shown in Figure 4.8.
The tree manages to create good splitting already at early levels. The first split is made on whether the pace is slow and the game relaxing: if so, awe experiences tend to be rare. Otherwise, challenging games seem to induce awe more frequently. Games that are not too challenging, though, still have good chances of inducing awe if the player enters the state of flow during the game, especially when they

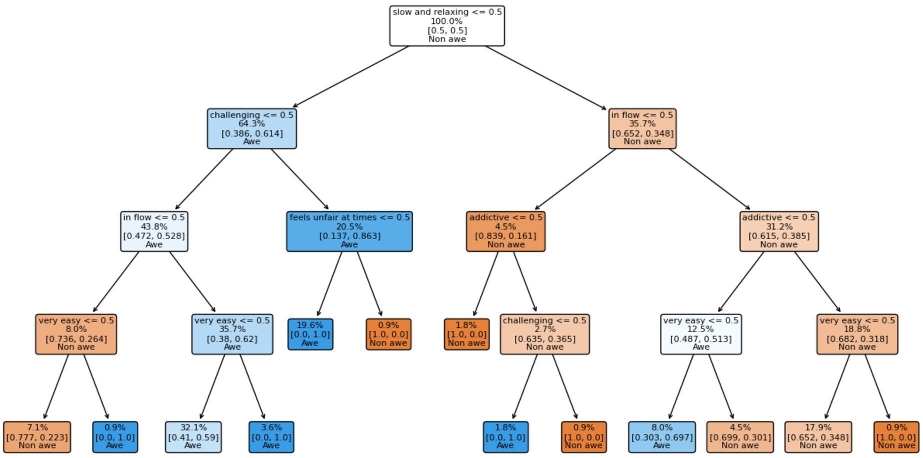

Figure 4.8: Decision tree for the 'pace and difficulty' dataframe

perceive the game as easy. On the other hand, it seems like addictive games tend to influence awe experiences negatively. The tree achieves a training accuracy of 75.9%, and a bit lower validation accuracy, of 67.9%.

Chapter 5

Discussion of results

From the previous analysis, some conclusions can be drawn about which video games features make them awe-inspiring. Knowing these features means having directions that can help prioritize certain aspects of a game already during the design phase, in order to improve the chances of provoking awe in players. The decision trees that have been trained on the individual dataframes have different performances that range from poor (in the case of those working with the ratings, graphics, and main character) to reasonably good (for example, the one analyzing pace and difficulty). This information should be taken into account when analyzing the selected features, in order to understand which results are reliable enough.
By looking at Figure 4.1, it can be seen that the most awe experiences occur when the soundtrack is rated very positively, with some exceptions. On the one hand, it seems that players that do not play very often can be awed even when the soundtrack is not 'perfect'. On the other side, awe seems to occur a bit regardless of the main character rating, in particular in frequent players when they appreciate the story of the game. Although some of these results align with the results of previous studies, especially regarding the soundtrack importance, others do not find an easy explanation based on what is known about awe; it has to be kept in mind, though, that this tree did not perform very well, so these results need to be confirmed with further research. However, a better understanding of how each game component influences awe experiences can be obtained by considering the other decision trees.
Figure 4.2 shows a link between simulation games and a lack of awe. This could be due to the fact that simulation games, although quite different from each other, tend to emulate aspects of our day-to-day life, while awe needs elements that are out of the ordinary experience - specifically, an individual has to feel like they are facing something larger than one's self - to be elicited. Few of them might, of course, still produce awe experiences: for example, flight simulators can easily include one of the best-known awe-provoking elements, the already mentioned scooping view of

vast scenery, and even with few samples, this case still shows up in the decision tree - not all simulation games actually fall into the ‘non-awe’ class. However, what can be understood from this tree is that awe experiences are uncommon with this (wide) category of games.
Considering Figure 4.3, there is a strong correlation between the graphics being described as ‘disturbing’ and awe. This does not come completely unexpected: it is to be remembered that awe has sometimes negative valence and that the feeling of threat is also a common trigger of the experience [35]. Disturbing graphics are likely to cause the players feelings that go from uneasiness to deep fear or terror; in the face of these emotions, a player might feel small, humbled, and might be compelled to try to make sense of what they are seeing - hence a possible source of awe.
Concerning the plot, the answers used in the second level of the tree in Figure 4.4 are not enough to draw conclusions about its significant attributes; more samples would be needed for a more precise analysis. However, the attribute used in the root node is more informative, as it alone allows for a decent split between awe and non-awe experiences. A possible explanation is that many subplots may build up the element of vastness, and as already mentioned, it is one of the core components of awe; as the player progresses through the game, all the subplots might relate to each other, creating the feeling that they are facing “the bigger picture" of the world they have been playing in.
Regarding the soundtrack features, irrelevant and unnoticeable music is unsurprisingly linked to a lack of awe; the majority of awe experiences happen when the soundtrack is not described as unnoticeable. These soundtracks might be characterized by any of the features that previous studies have identified as awe-inspiring: these include unexpected harmonies, sudden dynamics, the introduction of new voices, but also the change of sonic structure and distance, the use of crescendo, and major shifts in energy and volume, are commonly associated with the feeling of amazement, goosebumps, chills, and other physical manifestations of awe [36, 37, 38, 39, 40]. Music that includes such features would hardly ever be described as ‘unnoticeable’, and ‘pleasant’ might also not be the first adjective that comes to mind to describe them; instead, the adjectives that did come up in the answers to the survey were ‘epic’, ‘grandiose’, ‘emotional’, ‘chilling’, ‘wonderful’, which all seem to fit better the kind of music that inspires awe. These results, together with the ones shown in Figure 4.1, suggest what has already been explored in other studies about the strong connection between music and awe, highlighting the important role played by soundtracks in eliciting the emotion.
The decision tree associated with the main character dataframe does not seem to give very useful insights, as the first few nodes only classify single data points, except for the last one. The apparent correlation between perceived agency and awe might be related to an increased sense of freedom, that might even be missing

in the player's real life. However, the relationship that can be observed here might be of mere correlation, rather than causality. A sense of freedom is associated with awe and they can occur at the same time [41], but they might share common triggers, rather than trigger each other. The lack of literature supporting the claim that a sense of freedom might be an important factor in generating awe, together with the rather bad performance of the tree, makes it incautious to suggest that this might be the case; however, additional studies could be conducted to better investigate the relationship between the two. With the current data, it is difficult to conclude whether any of the features of the main character are able to induce awe; further research is needed to gain a better understanding of the subject.

Moving on to locations, Figure 4.7 shows that awe is sometimes perceived with games set either in modern cities or in a fantasy world; these two locations have the possibility to include known environmental triggers of awe. In the city setting, it appears that high buildings have the power to induce feelings of awe and smallness, usually associated with behavioral freezing [42]. In a fantasy world, instead, the source of awe could be found in the depiction of extraordinary natural environments, that often do not resemble what people are normally exposed to, and therefore trigger the process of accommodation. What counters expectations in this tree is that the space setting does not seem to induce awe experiences, even though it certainly has potential. This could be related to different factors: for example, the genre or the soundtrack, might both play a bigger role than the setting. For sure, a larger dataset could lead to finding a better explanation.

Finally, Figure 4.8 shows that awe does not commonly arise with slow-paced games; instead, it tends to be elicited either by challenging games or by games that get players in a state of flow, especially if they are perceived as easy by the players. The pairing between awe and flow does not come as a surprise: Chirico and Gaggioli [43] found a significant correlation between the two as self-transcendent experiences. The results from pace and difficulty in the tree could be related to the number of stimuli in the games: slow-paced and, especially, relaxing games might not give the player enough stimuli to make them perceive vastness, while a challenging game might achieve that a bit more easily. Another thing to keep in mind is that awe can be triggered, especially in people with individualistic cultural backgrounds, by personal accomplishments, which might be why challenging games can induce it [44]. Concerning addictive games, a hypothesis as to why they do not produce many awe experiences could be that they might absorb most of the player's mental energy to make them strongly engaged with the game, leaving less space for them to mentally process and take in something extraordinary.

To sum up, of all the examined video games features, the ones that seem to play a role in eliciting awe are: graphics that can shake players deeply, possibly causing fear; elaborated stories, rich in subplots; epic, majestic, chilling soundtracks; a modern city or fantasy world setting; a certain amount of challenge, enough to stimulate the players. However, further research is needed to confirm these results.

5.1 Limitations and future work

This work represents a starting point in finding how video games are able to elicit awe; however, even though there were some significant results, further effort is required to overcome some limitations. The first and most obvious one is the number of data points: with limited time and resources, the survey could not reach a very large audience, therefore the sample was too small for the intended analysis. Moreover, the dataset was not well balanced across different categories, from participants' demographics to awe experiences. This has an impact on the performance of all models, which could be largely improved with more data available. Alternatively, narrowing down the scope of the analysis could also lead to more significant results: for example, future studies could try to induce awe through one specific game or prototype, potentially one designed for the purpose, and directly observe participants playing it.

Another limitation to the validity of these results is that all the data was self-reported: in addition to the possible bias that some answers might be affected by, due to the time passing between the experience reported and the survey administration, each participant certainly had their own interpretation of the questions and the answers, possibly different from their intended one. Ideally, to minimize these issues, the data about the emotion of awe should be collected in person, by analyzing both physiological responses and self-reported experience, immediately after inducing the experience itself; collecting data in person also gives the chance to interviewers and participants to better clarify their points and ask for confirmation, reducing misunderstanding.

Another point to highlight is that this study focused on how the player described the game's attributes; descriptions are sort of broad, giving directions rather than pinpointing exactly what were the salient elements (e.g. graphics are disturbing, but it is not known whether that refers to hyper-realistic violence depiction, or to the looks of certain creatures, or to the details of the locations). The translation of those directions into concrete elements to include in the game can be addressed in future studies, for example by focusing on featuring specific elements and measuring their emotional effect on participants.

Chapter 6

Conclusions

This study had the goal to predict awe experiences with video games and to better understand them; after the initial collection of data, done through a survey, decision trees were trained with a varying degree of achieved accuracy, ranging from 50% to 75%. However, it is more interesting to look at the feature selection performed by the decision trees. Important characteristics of awe experiences were a highly-rated soundtrack and story, while unexpectedly the main character quality was negatively correlated with awe experiences. The simulation genre also seemed to induce awe very rarely, possibly because of the lack of out-of-the-ordinary stimuli. 'Disturbing' graphics seemed to be a recurring element in some of the awe experiences reported, most likely eliciting fear-related awe. An element of the story that seemed correlated with awe is the presence of several subplots in the game, which probably contribute to the perception of vastness, mentally challenging the player to comprehend it. Concerning soundtracks, it is not surprising that irrelevant or unnoticeable music was more often linked to non-awe, as awe-inducing music tends to be full of elements that make it feel 'epic'. The analysis of the main character yielded peculiar results, with the perceived agency seeming to lead to awe experiences; although a sense of freedom is often associated with awe, it is yet unclear what relationship exists between the two emotions. Locations that, unsurprisingly, were linked to awe were modern cities and fantasy worlds; both settings can indeed easily include established triggers of awe. While a slow-paced game might not often induce awe, it was found that presenting players with some challenge can lead to feeling awe, although this might be influenced by the cultural background of players and only apply to those who come from individualistic cultures.

Some of these results relate well to previous studies about awe, and what its common elicitors are - namely the results about soundtracks, locations, story, and partially the graphics - while others are a bit more difficult to explain, and might need to be double-checked with new data.

The study has some limitations: specifically, more data needs to be collected in order to clarify some of the results and to achieve better accuracy in the prediction of awe experiences - reflective of a good choice of the features that are actually significant for inducing awe. Narrowing down the scope of the analysis, maybe to a single game, might also give insights that have been missed in the general research. Moreover, awe would be more precisely identified by integrating different observation methods, like measurement of physiological responses and interviews, ideally to be conducted right after the experience. Finally, the intent of this work was not to focus on extremely precise elements, but rather to find directions for game designers to consider when thinking of inducing awe with their games.

Bibliography

[1] Dacher Keltner and Jonathan Haidt. «Approaching awe, a moral, spiritual, and aesthetic emotion». In: *Cognition and emotion* 17.2 (2003), pp. 297–314 (cit. on pp. 1, 3, 4, 7).

[2] JW Zhang and Dacher Keltner. «Awe and the natural environment». In: *Encyclopedia of mental health* 1 (2016), pp. 131–134 (cit. on pp. 1, 3).

[3] Leanne Pilgrim, J Ian Norris, and Jana Hackathorn. «Music is awesome: Influences of emotion, personality, and preference on experienced awe». In: *Journal of Consumer Behaviour* 16.5 (2017), pp. 442–451 (cit. on pp. 1, 3).

[4] Jessica J Luke. «"The Bloody Hell and Holy Cow Moment:" Feeling Awe in the Art Museum». In: *Curator: The Museum Journal* 64.1 (2021), pp. 41–55 (cit. on pp. 1, 3).

[5] Marianna Graziosi and David Yaden. «Interpersonal awe: Exploring the social domain of awe elicitors». In: *The Journal of Positive Psychology* 16.2 (2021), pp. 263–271 (cit. on p. 1).

[6] Denise Quesnel and Bernhard E Riecke. «Are you awed yet? How virtual reality gives us awe and goose bumps». In: *Frontiers in psychology* 9 (2018), p. 2158 (cit. on p. 1).

[7] Paul K Piff, Pia Dietze, Matthew Feinberg, Daniel M Stancato, and Dacher Keltner. «Awe, the small self, and prosocial behavior.» In: *Journal of personality and social psychology* 108.6 (2015), p. 883 (cit. on pp. 1, 4).

[8] Isabela Granic, Adam Lobel, and Rutger CME Engels. «The benefits of playing video games.» In: *American psychologist* 69.1 (2014), p. 66 (cit. on p. 1).

[9] Aaron Smuts. «Are video games art?» In: *Contemporary Aesthetics (Journal Archive)* 3.1 (2005), p. 6 (cit. on p. 1).

[10] Patrick Begley. «Empathy gaming'focuses on emotions and moral decisions'». In: *Sydney Morning Herald* (2014) (cit. on p. 1).

[11] Julia A Bopp, Jan B Vornhagen, and Elisa D Mekler. «"My Soul Got a Little Bit Cleaner" Art Experience in Videogames». In: *Proceedings of the ACM on Human-Computer Interaction* 5.CHI PLAY (2021), pp. 1–19 (cit. on p. 1).

[12] Alice Chirico and David B Yaden. «Awe: a self-transcendent and sometimes transformative emotion». In: *The function of emotions*. Springer, 2018, pp. 221–233 (cit. on pp. 3–5).

[13] David B Yaden, Scott Barry Kaufman, Elizabeth Hyde, Alice Chirico, Andrea Gaggioli, Jia Wei Zhang, and Dacher Keltner. «The development of the Awe Experience Scale (AWE-S): A multifactorial measure for a complex emotion». In: *The journal of positive psychology* 14.4 (2019), pp. 474–488 (cit. on pp. 3, 7).

[14] Craig L Anderson, Maria Monroy, and Dacher Keltner. «Awe in nature heals: Evidence from military veterans, at-risk youth, and college students.» In: *Emotion* 18.8 (2018), p. 1195 (cit. on p. 4).

[15] Mary Beth Oliver, Nicholas David Bowman, Julia K Woolley, Ryan Rogers, Brett I Sherrick, and Mun-Young Chung. «Video games as meaningful entertainment experiences.» In: *Psychology of popular media culture* 5.4 (2016), p. 390 (cit. on p. 4).

[16] Scott H Hemenover and Nicholas D Bowman. «Video games, emotion, and emotion regulation: Expanding the scope». In: *Annals of the International Communication Association* 42.2 (2018), pp. 125–143 (cit. on p. 4).

[17] Daniel Possler, Christoph Klimmt, and Nicholas D Bowman. «How Awe Affects Players' Entertainment Experiences Over Six Weeks of Playing». In: *International Conference on Entertainment Computing*. Springer. 2021, pp. 223–235 (cit. on p. 4).

[18] Daniel Possler, Christoph Klimmt, and Arthur A Raney. «Gaming is awesome! A theoretical model on cognitive demands and the elicitation of awe during video game play». In: *Video Games*. Routledge, 2018, pp. 74–91 (cit. on p. 4).

[19] Ken Fogelman and Chris Comber. «Surveys and sampling». In: *Research methods in educational leadership and management* (2002), pp. 93–107 (cit. on p. 6).

[20] Hendrik Müller, Aaron Sedley, and Elizabeth Ferrall-Nunge. «Survey research in HCI». In: *Ways of Knowing in HCI* (2014), pp. 229–266 (cit. on p. 7).

[21] Jon A Krosnick and Stanley Presser. «Question and questionnaire design. Handbook of Survey Research». In: *Education Emerald, London* (2010) (cit. on p. 7).

[22] Melanie Rudd, Kathleen D Vohs, and Jennifer Aaker. «Awe expands people's perception of time, alters decision making, and enhances well-being». In: *Psychological science* 23.10 (2012), pp. 1130–1136 (cit. on p. 7).

[23] Michelle N Shiota, Dacher Keltner, and Amanda Mossman. «The nature of awe: Elicitors, appraisals, and effects on self-concept». In: *Cognition and emotion* 21.5 (2007), pp. 944–963 (cit. on p. 7).

[24] *Intro to data structures*. `https://pandas.pydata.org/docs/user_guide/dsintro.html#dataframe`. Copyright 2008-2022 (cit. on p. 9).

[25] Julianna Delua. *Supervised vs. Unsupervised Learning: What's the Difference?* `https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning`. 2021 (cit. on p. 11).

[26] *Decision trees*. `https://scikit-learn.org/stable/modules/tree.html#tree`. Copyright 2007-2022 (cit. on pp. 17, 18, 20).

[27] F. Pedregosa et al. «Scikit-learn: Machine Learning in Python». In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830 (cit. on p. 18).

[28] Foster Provost and Tom Fawcett. *Data Science for Business: What you need to know about data mining and data-analytic thinking*. " O'Reilly Media, Inc.", 2013 (cit. on p. 18).

[29] Oded Z Maimon and Lior Rokach. *Data mining with decision trees: theory and applications*. Vol. 81. World scientific, 2014 (cit. on p. 18).

[30] Ignacio Icko. *Diagram showing overfitting of a classifier*. `https://commons.wikimedia.org/wiki/File:Overfitting.svg`. This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit `http://creativecommons.org/licenses/by/4.0/`. 2008 (cit. on p. 19).

[31] *Intro to data structures*. `https://pandas.pydata.org/docs/user_guide/dsintro.html#series`. Copyright 2008-2022 (cit. on p. 20).

[32] *Cross validation*. `https://scikit-learn.org/stable/modules/cross_validation.html#k-fold`. Copyright 2007-2022 (cit. on p. 20).

[33] *Feature selection*. `https://scikit-learn.org/stable/modules/feature_selection.html#select-from-model`. Copyright 2007-2022 (cit. on p. 23).

[34] *sklearn.tree.plot_tree*. `https://scikit-learn.org/stable/modules/generated/sklearn.tree.plot_tree.html#sklearn.tree.plot_tree`. Copyright 2007-2022 (cit. on p. 26).

[35] Amie M Gordon, Jennifer E Stellar, Craig L Anderson, Galen D McNeil, Daniel Loew, and Dacher Keltner. «The dark side of the sublime: Distinguishing a threat-based variant of awe.» In: *Journal of personality and social psychology* 113.2 (2017), p. 310 (cit. on p. 37).

[36] David Huron and Elizabeth Hellmuth Margulis. *Musical expectancy and thrills. Handbook of Music and Emotion: Theory, Research, Applications.* 2010 (cit. on p. 37).

[37] Oliver Grewe, Björn Katzur, Reinhard Kopiez, and Eckart Altenmüller. «Chills in different sensory domains: Frisson elicited by acoustical, visual, tactile and gustatory stimuli». In: *Psychology of Music* 39.2 (2011), pp. 220–239 (cit. on p. 37).

[38] Martin Guhn, Alfons Hamm, and Marcel Zentner. «Physiological and musico-acoustic correlates of the chill response». In: *Music Perception* 24.5 (2007), pp. 473–484 (cit. on p. 37).

[39] John A Sloboda. «Music structure and emotional response: Some empirical findings». In: *Psychology of music* 19.2 (1991), pp. 110–120 (cit. on p. 37).

[40] Peg Rawes. «Sonic envelopes». In: *The Senses and Society* 3.1 (2008), pp. 61–78 (cit. on p. 37).

[41] Lisbeth C Bethelmy and José A Corraliza. «Transcendence and sublime experience in nature: Awe and inspiring energy». In: *Frontiers in psychology* 10 (2019), p. 509 (cit. on p. 38).

[42] Yannick Joye and Siegfried Dewitte. «Up speeds you down. Awe-evoking monumental buildings trigger behavioral and perceived freezing». In: *Journal of Environmental Psychology* 47 (2016), pp. 112–125 (cit. on p. 38).

[43] Alice Chiricoa and Andrea Gaggioli. «The continuum of self-transcendence: Flow experience and the emotion of awe.» In: *Annual Review of CyberTherapy and Telemedicine* (2018) (cit. on p. 38).

[44] Summer Allen. *The science of awe.* John Templeton Foundation Greater Good Science, 2018 (cit. on p. 38).

Appendix A

Github repository

The full code used for this work can be found on GitHub at the following link:
`https://github.com/Aurora2701/MT_Code.git`

Appendix B

Survey

The next pages contain a printed version of the survey distributed for this study.

Awe experience with video games

Video games have been increasingly acclaimed for their potential to elicit intense feelings. One of them is awe, a complex emotion involving different elements, that can have either a positive or negative connotation.
If you think you might have felt awe while playing, answering this survey will help me gathering insights about the phenomenon for my master thesis. The survey will take about 10 minutes to complete.
Before diving deeper into the topic, I'd like to know a bit about you.

* Required

1. How old are you? Please write a number *

____

2. What is your gender? *

Mark only one oval.

- ( ) Male
- ( ) Female
- ( ) Prefer not to say
- ( ) Non-binary

3. How often do you play video games (on average)? *

Mark only one oval.

- ( ) less than 1 hour per week
- ( ) 1-2 hours per week
- ( ) 2-6 hours per week
- ( ) more than 6 hours per week

4. If you are a student, are your studies related to videogames?

Mark only one oval.

- Yes
- No
- Other: ______________

5. If you have a job, do you work in the videogame industry?

Mark only one oval.

- Yes
- No
- Other: ______________

Your experience

Now you are asked to think of a recent personal experience with video games when you believe you felt awe; because this is a complex emotion, you don't have to worry about any "rules" on how this experience should be. Instead, you can focus on whatever "awe" means to you.
If more than one experience comes to your mind, please focus only on one of them; if you want, you can always fill the survey again to give insights about the others.
Take your time to think; it might help to write down some notes about your experience, if you are finding it hard to recollect memories about it.
In this section, you have to express your agreement with a few sentences about your experience and how it made you feel.

6. I felt my perception of time change during the experience *

Mark only one oval.

	1	2	3	4	5	
Stro						Strongly agree

7. I felt like everything slowed down *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

8. I felt my sense of self become somehow smaller *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

9. I experienced a sense of oneness with all things *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

10. I felt humbled by the experience *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

11. I felt closely connected to humanity *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

12. I experienced something greater than myself *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

13. I had chills or goosebumps *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

14. I struggled to take in all that I was experiencing at once *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

15. I felt challenged to mentally process what I was experiencing *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

16. I think the experience had a positive impact on me *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

17. I felt fear or discomfort *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

18. I had a sense of peace of mind *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

19. I felt admiration towards the game developers *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

20. After the experience, I wanted to be a better person *

Mark only one oval.

	1	2	3	4	5	
Stro	◯	◯	◯	◯	◯	Strongly agree

21. Are there any other comments you wish to add about your feelings during or after the experience?

____

____

____

____

____

The video game

Thank you for telling me a bit about your experience; now, for this last section, I would like to know more about the game that triggered it.

22. What is the title of the game? *

____

23. What genre does the game belong to? Choose all that apply *

If you choose "Other...", please fill in all the genres that you think apply and are not mentioned in the other options, separated by a comma

Check all that apply.

- [ ] action (eg. platform, survival, stealth, shooter, fighting...)
- [ ] adventure (eg. text or graphic adventures, visual novels...)
- [ ] role-playing (eg. action RPG, MMORPG, roguelikes, monster tamer, JRPG...)
- [ ] simulation (eg. life simulation, vehicle simulation...)
- [ ] strategy (eg. turn-based strategy, tower defense, wargame, multiplayer online battle arena...)
- [ ] sports (eg. racing, sports-based fighting, competitive...)
- [ ] puzzle (eg. hidden objects, reveal the picture, tile matching, traditional puzzle...)
- [ ] action-adventure (eg. metroidvania)
- [ ] Other: ____________________

24. How would you rate the graphics? *

Mark only one oval.

	1	2	3	4	5	
Very	◯	◯	◯	◯	◯	Excellent

25. How would you describe the graphics of the game? Choose all that apply. *

If you choose "Other...", please add all the attributes that you are thinking about and are not mentioned in the other options, separated by a comma

Check all that apply.

- [ ] visually stunning
- [ ] realistic
- [ ] disturbing
- [ ] simple
- [ ] Other: ____________________

26. How would you rate the story (if not applicable skip this question)?

Mark only one oval.

	1	2	3	4	5	
Very	◯	◯	◯	◯	◯	Excellent

27. How would you describe the story of the game? Choose all that apply. (If the game did not have a story, select "not applicable") *

If you choose "Other...", please add all the attributes that you think apply and are not mentioned in the other options, separated by a comma

Check all that apply.

- [ ] complex
- [ ] rich in subplots
- [ ] includes plot-twists
- [ ] weak, contains plot holes
- [ ] captivating
- [ ] simple and easy to follow
- [ ] confusing
- [ ] boring
- [ ] not applicable
- [ ] Other: ____________________

28. How would you rate the soundtrack? *

Mark only one oval.

	1	2	3	4	5	
Very	◯	◯	◯	◯	◯	Excellent

29. How would you describe the soundtrack of the game? Choose all that apply *

If you choose "Other...", please write all the attributes that you are thinking of and are not mentioned in the other options, separated by a comma

Check all that apply.

- [ ] pleasant
- [ ] annoying
- [ ] on spot, perfect for the game
- [ ] chilling
- [ ] captured my attention while playing
- [ ] irrelevant, left me indifferent
- [ ] Other: ______________________

30. How did you like playing the main character (if applicable)?

Mark only one oval.

	1	2	3	4	5	
I hat	◯	◯	◯	◯	◯	I liked it a lot

31. Which of these items apply to your experience with the main character? You can select as many items as you want. (If the game did not involve controlling a character, select "not applicable") *

If your answer is not listed in the options, please select "Other..." and type your answer

Check all that apply.

- [ ] The character was customizable
- [ ] I could choose the character out of a pre-made set
- [ ] I felt represented by the main character
- [ ] I identified myself with the main character, to a certain extent
- [ ] I felt like my choices could affect the course of the game
- [ ] I was able to shape, at least partially, the main character's personality
- [ ] Not applicable
- [ ] The character was simple
- [ ] Other: ______________________

32. In what place(s) was the game set? Choose all the locations that you can remember from the game *

If the location is not among the options, select "Other..." and add all the locations that are missing, separated by a comma

Check all that apply.

- [ ] woods, forests
- [ ] desert
- [ ] sky
- [ ] sea, ocean
- [ ] space, spacecraft
- [ ] medieval towns
- [ ] modern city
- [ ] post-apocalyptic world
- [ ] fantasy world
- [ ] Other: ______________________

33. Which of these characteristics apply to the game? Choose all that apply *

If you choose "Other...", please fill in all the characteristics that you think apply and are not mentioned in the other options, separated by a comma

Check all that apply.

- [ ] the game was fast-paced and exciting
- [ ] the game was slow-paced and relaxing
- [ ] the game was very easy
- [ ] the game was challenging, sometimes too much
- [ ] the game kept me focused and absorbed
- [ ] the game was addictive
- [ ] the game was challenging at the right point
- [ ] Other: ______________________

34. When the experience occurred, were you playing using a virtual reality headset? *

Mark only one oval.

- Yes
- No

35. How would you describe the VR experience (if you replied "no" to the previous question, select "not applicable")? Choose all that apply *

If you choose "Other...", please fill in all the attributes that you think apply and are not mentioned in the other options, separated by a comma

Check all that apply.

- [ ] striking
- [ ] uncomfortable (even after practicing the controllers for a while)
- [ ] immersive
- [ ] wonderful
- [ ] indifferent
- [ ] distressing
- [ ] not applicable
- [ ] Other: ____________________

36. Is there anything else you would like to tell me about the game that was not covered by the previous questions?