2019-11-01

Recommending Games to Adults with Autism Spectrum Disorder(ASD) for Skill Enhancement Using Minecraft

Alisha Banskota
Brigham Young University

Recommending Games to Adults with Autism Spectrum Disorder (ASD) for Skill Enhancement Using *Minecraft*

Alisha Banskota

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Yiu-Kai Dennis Ng, Chair
Seth Holladay
Daniel Zappala

Department of Computer Science

Brigham Young University

ABSTRACT

Recommending Games to Adults with Autism Spectrum Disorder (ASD) for Skill Enhancement Using *Minecraft*

Alisha Banskota
Department of Computer Science, BYU
Master of Science

Autism spectrum disorder (ASD) is a long-standing mental condition characterized by hindered mental growth and development. In 2018, 168 out of 10,000 children are said to be affected with Autism in the USA. As these children move to adulthood, they have difficulty in communicating with others, expressing themselves, maintaining eye contact, developing a well-functioning motor skill or sensory sensitivity, and paying attention for longer period. Some of these abnormalities, however, can be gradually improved if they are treated appropriately during their adulthood. Studies have shown that people with ASD can enhance their social-interactive skills by playing video games. During the past decades, however, educational games have been primarily developed for autistic children, but not for autistic adults. We have developed a gaming and recommendation system that suggests therapeutic games to autistic adults which can improve their social-interactive skills. The gaming system maintains the entertainment value of the games, to make sure people are interested in playing them, whereas the recommendation system suggests appropriate games for autistic adults to play. Customizable games are designed and implemented in Minecraft such that each game focuses on enhancing different weakness areas in autistic adults based on games that the users have not explored in the past. The effectiveness of the gaming and recommendation system is backed up by an empirical study which shows that recommending therapeutic games can aid in the improvement of social-interactive skills of adults with ASD so that they can live a better life in the years to come.

Keywords: autism, adults, game development, recommendation system

ACKNOWLEDGMENTS

First of all, I would like to express a sincere gratitude towards my advisor Dr. Dennis Ng for the consistent support and guidance he provided me. Dr.Ng has always been very encouraging and has taught me the importance of hard work and punctuality. Dr.Ng has truly been a blessing with his support, both morally and technically. I would also like to thank my committee members Dr. Seth Holladay and Dr.Daniel Zappala for their support through my journey.

I would also like to thank the members of the advanced database lab. Without their efforts in development of the games and in providing me with their helping hands, the project would not have been a success.

Apart from that, I would also like to acknowledge the support given by my parents and my sister as I made the endeavour.Thank you.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Autism Spectrum Disorder (ASD), also known as autism spectrum condition (ASC), is a developmental disorder that affects social, behavioral, and communication skills of an individual. It is called a "spectrum" disorder because the disorder may vary from mild to severe and appear differently in different individuals. ASD encapsulates a number of neurodevelopmental disorders including autism, Aspergers syndrome (AS), and pervasive developmental disorder not otherwise specified. Autism is considered the more severe in the spectrum while AS is considered lighter [27]. Individuals with ASD face challenges in social engagement and in developing appropriate relationships with their peers at different deficiency levels [16]. There is a high demand for offering treatments for people with ASD, since according to Centers for Disease Control and Prevention (CDC) Autism and Developmental Disabilities Monitoring Network, about 1 in 59 children are identified with ASD [1]. Every year, 50,000 teens are entering into adulthood with ASD [1] who are prevalent in all racial, ethnic, and socioeconomic groups and autism is about four times more common in boys than in girls.

People with ASD face different challenges that could lie in cognitive, behavioral, motor sensory, or medical areas [25]. Despite the limitation in social or day-to-day activity, individuals with ASD are highly interested in video games as shown in different studies [14]. Children and adolescents with ASD spend more than two hours a day playing video games, which is more than by typically developing children or children with other disabilities [15]. Not just for children, adults with ASD have also reported to be engaged in video games and these video games are said to have positive impact on them as the adults find video games to be entertaining, stress relieving, and creative [15].

Regular video games are neither targeted to people with ASD, nor do they focus on their weakness areas. For this reason, regular video games are not designed especially for adults with autism in mind. However, games have been previously developed for children with ASD to enhance their social skills. There are specific games [9] developed that are both fun and acting as a therapeutic agent for children with ASD. These systems, however, are not suitable for adults with ASD as these games are too primitive for adults and they do not enjoy or learn from them. Also, adults with ASD face different challenges than children with ASD which means that there is a demand to design therapeutic games especially for adults with ASD. Recent studies show that adults with ASD are also preoccupied with screen-based media, such as television and video games [14]. In fact, 64.2% of young adults with ASD spend their free time using non-social, screen-based media [14].

There is a need to develop a gaming system played by adults with ASD to enhance their social-interactive skills in order to connect autistic adults with each other and with typically developed adults. Development of social-interactive skills can help adults with ASD become independent and lead a fulfilling life [28]. These social-interactive skills assist them in improving their daily living skills i.e, day-to-day skills required to live independently such as, transportation, budgeting and lifelong skills such as getting education [28]. Unfortunately, very little has been explored in the area of video games for social-skill enhancement of autistic adults.

Developing well-designed therapeutic games for adults with autism is essential, since they can provide entertainment value to players and better engage autistic adults to learn social skills. If players are not attracted to a game, they will not spend time playing the game. Games that can be both entertaining and educational give rise to the concept of ‘edutainment [24]. Such edutainment games first started appearing in the early 80’s [6] and today they are estimated to be an $8.4 billion market in the USA and include games for both children and adults [10].The effectiveness of idea of 'edutainment' can be verified through the case study of [12] where a gamified collaboration system is used on an introductory programming course and the system encourages students to participate and contribute more towards the course.

We have developed a gaming and recommendation system with 'edutainment' values. Among a myriad of options available as base to develop our games, we have chosen Minecraft because people with ASD enjoy Minecraft and it is quick and easy to develop games in Minecraft. Apart from that, it also allows the player to be creative, is captivating, and is available across various computer platforms, such as PC or Mac.

Minecraft, which is a game environment that is quick to develop, is a sandbox video game that allows multiple players to create their own world or build structures inside a preexisting world. Apart from that, players can also collect resources, explore the world, and even destroy blocks in the world. Minecraft gameplay has five distinctive modes, and each mode offers different privileges and access control to the user. Multiple players can be a part of the game even remotely and Minecraft provides plugins for them to interact with each other through chat or voice. We have used these features of Minecraft to create games that focus building on social-interactive skills lacked by adults with ASD.

We created therapeutic games related to different areas of weakness faced by adults with ASD and recommended the most appropriate ones to each player based on his/her weakness. In order for our recommendation system to work, we relied on a finite set of areas of weakness established by ScenicView Academy, a school in Provo, Utah for autistic adults. ScenicView has published a booklet called "Bridges we build – the art of making friends" [28], which is a guide to connect and maintain socialized relations with others. The booklet includes information about different steps to be taken to maintain social relationships, why they are necessary, and how those steps can be taken. For example, they talk about why people should connect with each other, how to find friends, how to seek positive attention, etc. Based on these principles, we have derived the most prominent challenges faced by adults with ASD which comprises of *developing audio communication skills, recognizing facial expressions, maintaining eye contact, showing empathy*, and *engaging in speech therapy.* We realize that the nature of these challenges and their severity may differ person to person.

Since the degree of severity of ASD is subjective to an individual adult, we have developed a recommender system taking this fact into account. The system determines which of the skills need most improvement and which of our games each adult should use. We start with studying the weakness areas of autistic adults based on their *gaming profiles* and a *questionnaire* containing questions that address to different weakness areas. Hereafter, we suggest the most appropriate games to them so that the adults could work on the areas of deficiency.

To verify the performance of our recommendation system in terms of improving the social-interactive skills of adults with autism, we have conducted an empirical study on a group of students at ScenicView. We observed the changes that occurred among these students during the process. The experimental results reflect the changes in social behavior observed after the participants played the recommended therapeutic games. With the results extracted from the empirical study, we conclude that our recommended games do play a role in changing their social-interactive skills. As the ultimate design goal of our recommender system is to enhance the social-interactive skills of adults with ASD, we claim that the design goal of our system has been met.

Chapter 2

Related Work

A large number of gaming systems have been developed in the past, since 97% of teens between the age of 12-17 play some kind of video games [5]. Hilary Smith [18] talks about the positive impact of video games on social skills, flexibility and motor skills of children with special needs. Parents of children with ASD generally support video game play which encourages the use of video games in the ASD community [21]. There are some therapeutic gaming systems targeting specifically towards the community. The use of Serious Games, which is a video game with simulations to develop skills, has been increasing in the field of ASD [30] and has been effective in improving different skills such as mathematical skills, decreasing anxiety, etc. There are games developed specifically for children such as Go Go Games [9] for elementary school children with ASD to teach them "multiple cue responding", Autcraft [26], which is a Minecraft server especially developed for children with ASD, and the Flexibility Learning Project (FLOW) [22], which works on rigidity of children with ASD. The authors of [20] discuss a review on development and research in serious games for children with Autism. Their study shows that games are very effective for therapeutic purposes such as improving social skills and behavior, working on co-ordination skills etc. as well as for educational purposes such as first aid learning, concept of money etc. These design and researches in the field of ASD, however, are concentrated mostly for children and thus, they mainly focus on simple activities and life events. Adults who are already aware of those activities would not be interested in playing those games. The authors of [8] study about serious games to teach social interactions and emotions to individuals with

ASD. Through multiple platforms, they find a total of 31 serious games out of which only 2 projects called *Computer-based program* and *Lets face it* are designed for young adults and adolescents. The results of majority of the 31 projects show that individuals with ASD showed significant improvement through those projects. This strengthens the necessity of developing more video games specially targeted for adults with ASD.

The authors of [15] have conducted a thorough study of perspective of adults with ASD on video games. The study includes a survey with 58 adults with ASD who were asked four primary questions about the video games that they play. The survey offers an insight on the views of adults with ASD towards video games. This insight is an affirmation to the vision we obtain from ScenicView Academy [28] for skill enhancement of autistic adults. Along with that, they also categorize games into mutually exclusive genres which include action-adventure, role playing, music, fighting, puzzle, sandbox, etc. This concept of genre is similar to the concept of labels and key phrases we use in our gaming and recommendation system to characterize games for autistic adults.

Both gaming and non-gaming approaches have been used in the past to enhance the skills of autistic adults. The authors of [2] address the usefulness of Covert Audio Coaching (CAC) in which the high school students with ASD were given feedback as they were performing the task of making photocopies. The idea is to examine the impact of performance feedback on their skills. Our gaming approach, however, is to develop such a system whose activities are not just instructive but enjoyable to the adults, and games fit perfectly into that category. The motivation behind developing an enjoyable approach is to ensure that people are interested and engaged while using our system.

There are also a few gaming systems, designed especially for adult's cognitive or behavioral disabilities. The authors of [29] design two games for people with cognitive disabilities. The first game is developed to help a worker with cognitive disability to fit into a new work environment and the second game is created to train those adults on social and self-autonomy skills. Based on the case study with these games, they discuss the potential improvements that could be made for adults with cognitive disabilities. The authors of [19] also recommend social-interactive games to adults with ASD. The authors generate both a short-head (familiar to the user) and a long-tail (unfamiliar to the user) game recommendation using features of video games such as the ratings, topic relevance, complexity of games to make recommendation but they do not take user profile into consideration for that. They separately show that both of their approaches perform better than other existing approaches. So we came up with a recommender that relies both on the labels of video games and as well as user profiles. Each set of labels correspond to the skills of the autistic adults which they can work on.

The authors of [11] argue that games with 'edutainment' purpose have failed because they are incapable of increasing the endogenous value of a game without compromising the fun aspect. Keeping this in mind, we have attempted to develop games that have both endogenous value of enhancing the skills lacked by autistic adults as well as entertainment values to keep the players engaged.

Most approaches that have developed or studied the impact of video games in people with cognitive disabilities use the pretest posttest design where measurements are taken before and after the usage of the approaches [23]. We also use a similar approach for our experimental study where we observe the changes in users before and after they play our therapeutic games with the help of the questionnaire.

Chapter 3

Our Gaming System

We divide the design of our gaming and recommendation system for autistic adults into two phases. The first phase is the development of therapeutic games, which is presented in this chapter. The second phase, which is the recommendation system that suggests to the users the most appropriate games according to the social-interactive skills they lack, will be introduced in Chapter 4.

3.1 Game Development

We have designed and implemented different therapeutic games in Minecraft such that they cater to at least one of the five weakness areas: (i) developing audio communication skills, (ii) engaging in speech therapy, (iii) recognizing facial expression, (iv) showing empathy, and (v) maintaining eye contact. These therapeutic games, are effective as they have fun aspect and they are designed to address the weakness areas of their players.

3.1.1 Developing Audio Communication Skills

Adults with ASD face challenges in maintaining communication, listening to others or understanding other's perspectives. As a result, they have difficulty in forming relationships, expressing themselves to others, etc. Along with others, we have developed a *communication room* game in which two or more players come together and have a regular conversation with one another. At the end of a conversation, the players are asked to take a quiz about each other. Based on the answer they give, they either pass the test if they answer correctly

(a) Screenshot of communication quiz game for developing audio communication skills

(b) Roleplay game for engaging in speech therapy

Figure 3.1: Games for developing audio communication skills and speech therapy

about their partner or are suggested to spend more time talking to each other. The quiz is designed such that the players cannot pass the test without conversing with one another which ensures the effectiveness of the game. The game has hints to help them choose a conversation topic or keep the conversation going. This enforces people who are not very effective in communicating to find new people in the game, initiate a conversation, learn how to keep a conversation going, and make friends. Figure 3.1(a) shows the screenshot of a game for developing audio communication skills.

3.1.2 Engaging in Speech Therapy

Adults with ASD can have problems paying attention to sounds of others, decipher them, or produce sounds themselves. These are some challenges that are unique to people with Autism, i.e., pronoun reversal, repeating what someone has said, and producing less sophisticated language than they comprehend. Challenges in speech again impedes them from expressing themselves or understanding others. We have created a *multiplayer roleplay* game, one of the games to encourage autistic adults engaging in speech therapy, in which the world is going to collapse because of zombie attack and only two people among all can be saved. Each player chooses a role available and there is a judge who is invigilating the game. Each player has to convince the judge why (s)he has to be saved. This game requires players to clearly understand what other players are saying and at the same time articulate their points with clarity. The judge in the game evaluates the effectiveness of the game and the performance of the players. Figure 3.1(b) shows screenshot of the game.

3.1.3 Recognizing Facial Expression

Recognizing facial expression is very crucial for social interaction, and studies have shown that in typical development, even an infant can distinguish facial expressions. However, for adults with ASD, understanding the situation or what the other person is saying through their expressions can be confusing. Our facial recognition game which caters to identifying facial expressions in Minecraft, asks the player to determine facial expressions in a picture. The users can reach the end of the game and get rewards only if they complete it which ensures that they get a good practice with facial expressions. This kind of game is a good exercise for the users to learn what basic expressions, such as happy, sad, angry, surprised, shocked, tired, etc., are like. Figure 3.2(a) shows the screenshot of a facial expression recognition game.

3.1.4 Showing Empathy

Some autistic adults have problems showing empathy or expressing how they feel in a sensitive situation which is a major underlying cause behind deficiencies in social-interaction or relationship building. To address this issue of showing compassion, we have developed different games where players communicate with a fictional character. In this type of game, the character interacts with the player and shares his thoughts.

In showing empathy games, each player has to choose the most appropriate response based on how the character is feeling out of a multiple choice of options given to them. If the player chooses a wrong response, the character reacts accordingly, which shows the player how the responses could be insensitive. When the user plays till the end of the game, we

(a) Jim's World game for recognizing facial expression (b) Screenshot of game for showing empathy

Figure 3.2: Games for recognizing facial expression and showing empathy

can assume that the game is fun and effective. Figure 3.2(b) shows the screenshot of the game with the character called *Victor*.

3.1.5 Maintaining Eye Contact

Maintaining eye contact is essential to know if the other person involved is interested in the conversation. For some adults with ASD, maintaining eye contact during a conversation makes the situation even more stressful as they have to talk and look at someone at the same time. We have designed multiplayer games that encourage players to maintain eye contact with the other player. One of these games is a multi-stage game in which there is a dungeon that the players have to pass through. Each player is prompted to talk about their favorites, such as movies, games, food, etc. If both players maintain eye contact during the conversation, they are moved through the dungeon. We have the game such that two people are maintaining eye contact if they are connected for 4 to 5 seconds [4]. The game first detects if the co-ordinates of two players are complimentary[1] and determines if the co-

[1]If two players are facing each other, the co-ordinates of one player is equivalent to 180 degree rotation of the other player.

Table 3.1: Games developed for our gaming and recommendation system

Game	Weakness Area	Description
Communication Quiz	Developing Audio Communication Skills	Two players talk to one another and take a quiz about each other
Dungeon Journey	Maintaining Eye Contact	Players maintain eye contact while communicating with one another to get out of dungeon
Jim's World	Recognizing Facial Expression	Players detect how Jim is feeling based on his expressions
Make New Friends	Engaging in Speech Therapy, Recognizing Facial Expression	Player listens to voice recordings to choose the expression of the displayed image
Role Play Debate	Engaging in Speech Therapy	Players take up a role and argue about why they should be saved from apocalypse
Survival World	Developing Audio Communication Skills	Players build a shelter to protect themselves from zombies
T-Shaped Bridge	Developing Audio Communication Skills	Players help one another to build a bridge to escape
Team Tunnel	Developing Audio Communication Skills, Showing Empathy	Players help one another get out a tunnel
Voice	Showing Empathy	Player listens to a character called Victor and give Victor the right advice
Zombie Maze	Developing Audio Communication Skills	Players help each other get out of a maze that has zombies

ordinates remain the same for 5 seconds or more. If the co-ordinates of a player changes before 5 seconds, we assume that the player is distracted. The game then prompts the players to maintain eye contact in order to continue.

3.2 The Developed Therapeutic Games

Along with the therapeutic games mentioned in previous sections, there are other games that we have developed. The exhaustive list of games that have been designed and implemented, along with the areas they relate to, are shown in Table 3.1 .

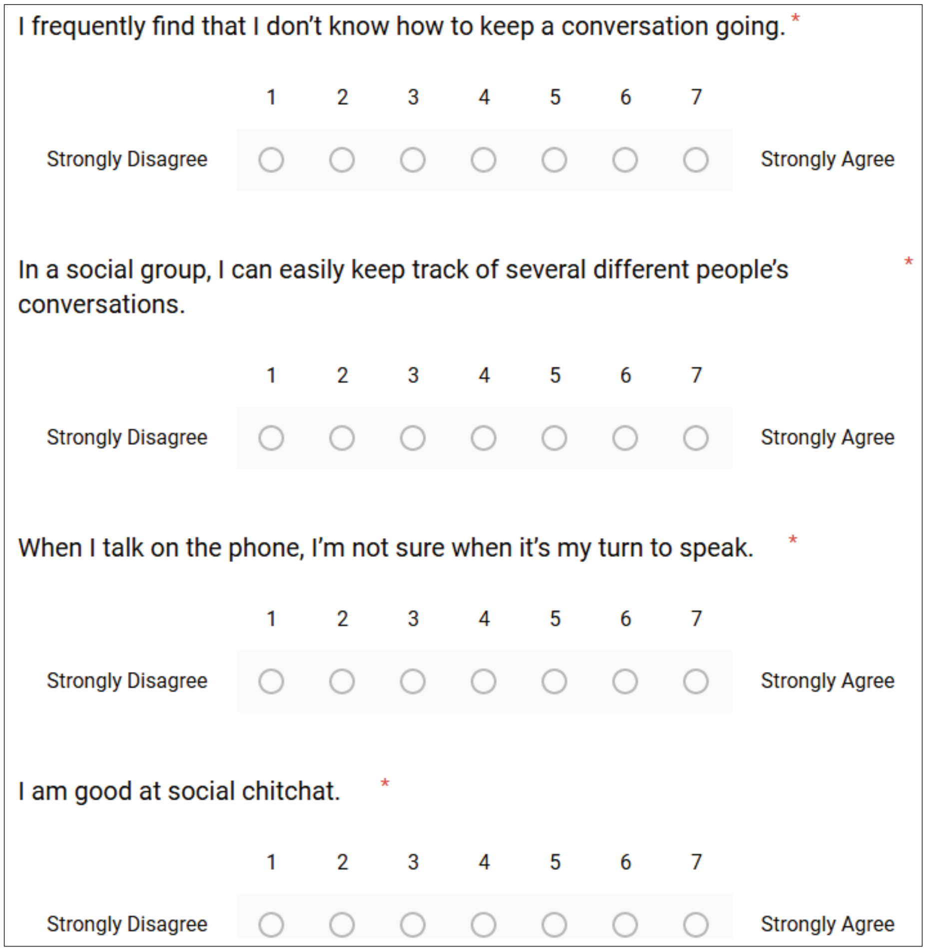
Chapter 4

The Basic Design of Our Game Recommendation System

Our recommender system suggests games to users based on their weakness areas so that they can enhance their social-interactive skills in the respective areas. Since it is difficult to obtain user's clinical data as it is confidential, we infer the weakness areas faced by autistic adults through the VideoGameGeek games they have played previously. VideoGameGeek archives thousands of games and provides different metadata such as the themes of each game, number of players in that game, platform in which the game is played, description of the game, etc. These games are *not* therapeutic for autistic adults, since they are only developed for entertainment or educational purpose. On the other hand, therapeutic games that we have developed in Minecraft are designed to address different weakness areas prevalent in autistic adults. We use *metadata* about games from VideoGameGeek to rank our Minecraft games for recommendation as discussed in the later sections. We begin the recommendation process by using a *questionnaire* and then analyzing the answers to its questions for actual game recommendation.

4.1 Using Questionnaire to Determine (the Effect of) Therapeutic Games Recommended to Users

To recommend the most appropriate therapeutic game developed by us to each user, we create a questionnaire and ask the user to it out so that we can study (the changes in terms of) the deficiency areas of the user. Questions in the questionnaire co-relate to each of the *five* weakness areas and the answers given to these questions indicate the severity of



I frequently find that I don't know how to keep a conversation going. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

In a social group, I can easily keep track of several different people's conversations. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

When I talk on the phone, I'm not sure when it's my turn to speak. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

I am good at social chitchat. *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

Figure 4.1: Sample questions in the questionnaire (* means all questions are mandatory)

deficiency of the user in those areas. (A number of sample questions in the questionnaire is shown in Figure 4.1.) The users choose if they 'Strongly Disagree', 'Disagree', 'Slightly Disagree', 'Neither Agree Nor Disagree', 'Slightly Agree', 'Agree', and 'Strongly Agree' to each question. The exhaustive list of questions in the questionnaire with their corresponding weakness areas is shown in Table 4.1.

Prior to recommending any therapeutic games to autistic adults, we require each user to answer the questions in the questionnaire. For each question, a score between 1 to 7 inclusively is assigned as the answer by the user depending upon the positivity connotation of each answer such that if the answer indicates the absence of a weakness area, it is assigned a value closer to the upper half of the range of values and if it indicates the presence of a weakness area, it is assigned a value closer to the lower half instead. Since the questions are associated with the five weakness areas, after the user has filled out the questionnaire, we take the *average score* of the answers to the questions belonged to each weakness area. We take the average because each weakness area is addressed by multiple questions and we would like to obtain the mean overall score for each weakness area. The average score computed for each weakness area yields the *baseline score*. We set **5.5** as a *threshold* so that a value 5.5 or higher indicates positivity, i.e., absence of weakness, and below 5.5 implies negativity. The threshold value 5.5 is used to determine whether to continue recommending games in a particular weakness area to be played by a user to enhance his social-interactive skill or to make the conclusion that the user has no deficiency in the corresponding weakness area (anymore).

In order to find out all the weakness areas a user is struggling with, we ask the user to fill out the questionnaire before playing any of the therapeutic games. The *baseline scores* of all the weakness areas computed for a user at this stage is called the *initial baseline scores* of the user. The index of weakness areas in the initial baseline scores is the same for all the users. If the initial baseline score corresponding to a weakness area indicates the presence of a particular weakness, we recommend therapeutic games in that area. However, since there is a number of therapeutic games that can be recommended in each of these areas, we develop a *ranking strategy* of those games to decide the sequence in which we recommend them to the user.

4.2 Ranking Therapeutic Games

We have developed a number of therapeutic games in each of the weakness areas and rank those therapeutic games based on their relevance towards each user. For that, we rely on the top-10 VideoGameGeek games that the users have enjoyed playing in the past. Based on this information, we then rank therapeutic games belonged to each weakness area using *attribute* *values* of the VideoGameGeek games. Attributes of VideoGameGeek games considered by our recommender are shown in Table 4.2.

4.3 Attributes of VideoGameGeek games

We calculate the overlap of the attribute values between the VideoGameGeek games that the user has previously played with the attribute values of the therapeutic games to rank the latter in order to generate a sequence of recommended games for a particular weakness area. The overlap of attributes is calculated in order to ensure that we rank the therapeutic games based on the preference of games that users have played in the past so they enjoy playing the therapeutic games. The therapeutic game, whose attribute values are *closest* to the attribute values of the VideoGameGeek games that the user has played previously, is ranked the *highest* followed by the second closest game and so on.

4.3.1 Themes

A theme of a VideoGameGeeek game is the *category* to which the game belongs. There are 70 unique themes among all the games in VideoGameGeek (shown in Figure 4.2) and each game can have one or more themes associated to it. We manually assign one or more of these themes to each therapeutic game as well. In order to rank the therapeutic games to be recommended based on their themes, we create a *co-relation matrix* among *themes* to determine their *frequencies of co-occurrence*. The therapeutic game whose themes have the *highest co-relation values* with the themes of the VideoGameGeek games played previously by the user partially determines its ranking among other therapeutic games.

To compute the co-relation matrix, we first create a matrix of size 70×70 as there are 70 distinct VideoGameGeek themes and initialize each entry as zero. For each theme of a VideoGameGeek game, we create an index to be used in the matrix and count the number of games with that specific theme. The matrix is filled with the number of times each theme

abstract, adult, agriculture, aliens, amusement park, animals, anime, art, board game, business, card game, cartoon, children, comedy, comics, cooking, crime, cyberpunk, dancing, detective, driving, Egyptian, fairy tale, fantasy, fishing, flight, folklore, history, horror, hunting, martial arts, medical, medieval, military, minigame compilation, mining, monster collecting, moral choices, morbid, movie and book and tv show, music, mythology, nature, ninja, occult, ocean, other, physics, pirates, politics, public transport, puzzle, relationship, robot, samurai, school, science and technology, science fiction, space, sports, spy, steampunk, superhero, surreal, talking animals, Viking, war, western, words, zombies

Figure 4.2: Themes extracted from VideoGameGeek

is co-occurrent with one another in the same game. In other words, it keeps track of how many times *two* themes appear in the same game, which is computed as follows:

$$CoFreq(V, W) = \frac{P_{V,W}}{P_{V,V} + P_{W,W} - P_{V,W}} \tag{4.1}$$

where $CoFreq(V, W)$ is the *co-occurrence frequency* between themes V and W among all the VideoGameGeek games, $P_{V,W}$ is the *frequency of co-occurrence* of themes V and W together in the VideoGameGeek games, $P_{V,V}$ and $P_{W,W}$ are the number of times themes V and W appear individually among all the games in VideoGameGeek, respectively.

Using this process, we obtain the matrix, which is partially shown in Figure 4.3, where each value is a co-relation between two themes and the range of each value is between 0 and 1. The diagonal represents co-relation of the same theme, which is always 1.

After we have calculated the co-occurrence frequency between two themes, we can determine the *category similarity score* among each therapeutic game and the VideoGameGeek games specified in a user profile using Equation 4.2.

$$\begin{aligned} CTS(TG) &= \sum_{j \in VG} \frac{\sum_{(V,W) \in S(j,TG)} CoFreq(V, W)}{|TG| \times |VG_j|} \\ S(j, TG) &= \sum_{i \in VG_j} \sum_{k \in TG} (i, k) \end{aligned} \tag{4.2}$$

where $CTS(TG)$ is the *theme similarity score* of a particular therapeutic game denoted TG, $|TG|$ is the number of themes in TG, $|VG_j|$ is the number of themes in the j^{th}

	abstract	adult	agriculture	aliens	amusement park	animals	anime	art	board game	business
abstract	1	0	0	0	0	0.008929	0.008403	0.014493	0.008403	0
adult	0	1	0	0	0	0	0	0	0	0.010417
agriculture	0	0	1	0	0	0	0	0	0	0
aliens	0	0	0	1	0	0	0	0	0.016949	0
amusement park	0	0	0	0	1	0	0	0	0	0
animals	0.008929	0	0	0	0	1	0	0	0	0
anime	0.008403	0	0	0	0	0	1	0	0	0
art	0.014493	0	0	0	0	0	0	1	0	0
board game	0.008403	0	0	0.016949	0	0	0	0	1	0.009174
business	0	0.010417	0	0	0	0	0	0	0.009174	1

Figure 4.3: A portion of the co-relation matrix of themes generated using VideoGameGeek games

VideoGameGeek game, denoted VG_j, in the user profile, $S(j, TG)$ is the set of any two themes created using the cross product of themes in the j^{th} VideoGameGeek game in a user profile and TG, $CoFreq(V, W)$ is the *co-occurrence frequency* between two themes V and W defined in Equation 4.1.

4.3.2 Topic Analysis

There is a *textual description* for each game in VideoGameGeek explaining what the game is and we have created our own descriptions for the therapeutic games as well. Different *topics* can be assigned to these descriptions, and determining games with the *same* topics is a feature adapted by us in *ranking* our therapeutic games. In order to collectively rank the therapeutic games to be recommended to the users, we consider topics of the therapeutic games and the VideoGameGeek games. The therapeutic games whose descriptions are *closely related* to that of the VideoGameGeek games are ranked *higher*. We train a **Latent Dirichlet Allocation (LDA)** model to generate a list of unique topics formed with the keywords that are best associated with the game descriptions. Hereafter, we assign the topics that match our therapeutic games with the VideoGameGeek games played by the users in the past to develop a ranking based on topics.

Latent Dirichlet Allocation (LDA)

LDA is a generative probabilistic model that is mostly used in *topic modeling*. Topic modeling refers to identifying the appropriate topics that best describe a set of documents. These topics, which are not pre-assigned and are created using the modeling technique, are called *latent* or *hidden* topics. Each topic in LDA is a *collection of keywords* and the central idea is to match each document to a topic such that the keywords in the document are reflected by the keywords belonging to the topic.

The LDA model consists of two sets of matrices. The first matrix determines the probability of selecting a particular *word* given a *topic* and the second one finds out the probability of selecting a particular *topic* given a *document*. Based on these matrices, the model determines the probability for each of the topics given a document. During the training, LDA determines the probability of a word w given a topic z, i.e., $P(w|z)$, and the probability of a topic z given a document d, i.e., $P(z|d)$. We apply Gibb's sampling to train the LDA model. The Gibb's sampling assigns each word w present in the set of documents to a separate cluster, i.e., the cluster has only 'w' as its element. Hereafter, for each of the documents d in the set, if the cluster contains a word mentioned in the document, the document is included in the cluster. During this process, Gibb's sampling estimates $P(w_i|z_j)$ and $P(z_j|d_k)$ by iterating over each word w_i in each document d_k and assigns a cluster for w_i based on the probability, which is calculated as follows:

$$P(z_i = j|w_i, d_k, z_{-i}) \propto \frac{C_{WZ}(w_i, j) + \beta}{\sum_w C_{WZ}(w_i, j) + W\beta} \times \frac{C_{DZ}(d_k, j) + \alpha}{\sum_z C_{DZ}(d_k, z) + Z\alpha} \tag{4.3}$$

where $P(z_i = j|w_i, d_k, z_{-i})$ is the probability in which z_i is the *topic* assigned to w_i and z_{-i} denotes all topic-and-word and document-and-topic assignments excluding the current assignment z_i for word w_i, the first multiplicative factor in the equation computes $P(w_i|z_i =$

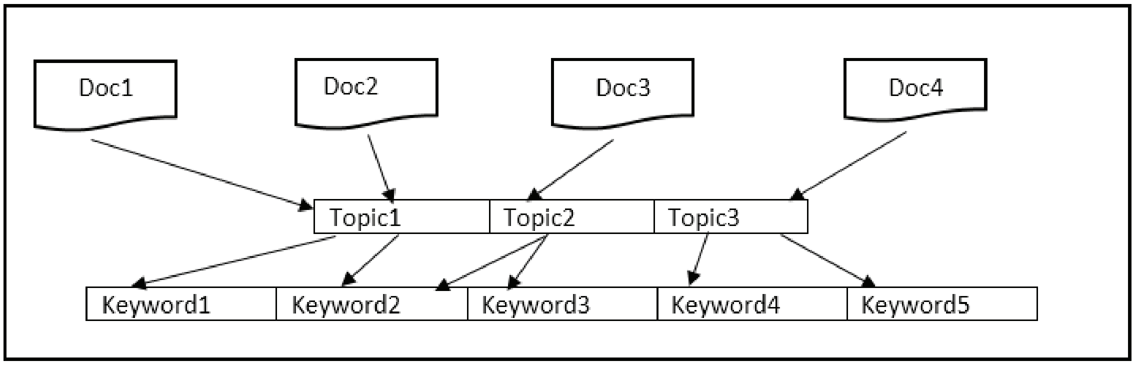


Figure 4.4: Illustration of topic modeling through LDA which shows documents associated with topics which as in turn associated with keywords

j), i.e., the probability of a word given a topic, and the second factor calculates $P(z_i = j|d_k)$, i.e., the probability of a topic given a document. C_{WZ} and C_{DZ} are the count of topic-and-word and topic-and-document, respectively, such that $C_{WZ}(w_i, j)$ is the number of times the i^{th} word, denoted w_i, in the document d_k is assigned to topic j, $C_{DZ}(d_k, j)$ is the number of distinct words in d_k that are assigned to topic j. Z is the total number of topics, W is the total number of distinct words in the set of documents, and α and β are *smoothing parameters.*

Documents could be modeled by words themselves, instead of topics; but for a large collection of documents, the number of words to represent the documents would be very large, and each of the documents has to be represented by a large number of words which is not feasible. Therefore, imaginary topics are created using representative keywords which are then used to model the documents. An illustration of the process is shown in Figure 4.4. In Figure 4.4, all the documents (Doc1 - Doc4) are associated to one of the three topics (Topic1 - Topic3) and each of those topics are related to two keywords (Keyword1 - Keyword5). Each keyword in a topic has a weight associated with it that reflects how important that keyword is for the topic. In order to assign a topic to a document, LDA assigns probability to each of the topics for the document and the topic with highest probability is chosen as the topic for the document.

Portal 2 will continue to challenge the player by solving puzzles in test chambers within the Aperture Science Enrichment Center using the portal gun (the Aperture Science Hand-held Portal Device), a device that can create two portals connecting two surfaces across space. Players, as the silent protagonist Chell from Portal, solve puzzles by using these portals to move unconventionally between rooms or to use the ability to fling objects or themselves across a distance. The functionality of the gun has not changed between the games, but within Portal 2, players can take advantage of the bleeding of other physical effects through the portals.

Figure 4.5: The description of the VideoGameGeek game 'Portal 2'

The Training Process

To train the LDA using game descriptions, we extract the description of around 2,000 games from VideoGameGeek and preprocess the descriptions of each game. We first remove the stop words, stem the words, and construct the trained LDA on the pre-processed description of VideoGameGeek games based on the 100 keywords used to distinguish each topic along with their weights. We consider LDA for different number of topics starting from 20 to 50 and manually study the results of LDA for different number of topics. Eventually, we conclude that twenty-seven is the optimal number of topics as we were able to associate the descriptions of all the therapeutic games to the twenty-seven topics, and the topic assigned to each of the descriptions captured the important keywords of the description.

Topic Modeling on Game Descriptions

In order to train an LDA and develop a partial ranking for the therapeutic games, we represent the descriptions of the games as documents. (An example of description of the game 'Portal 2' is shown in Figure 4.5). Hereafter, we run LDA on each document, which is the description of a therapeutic game, to obtain the topic that best matches the description and compare the chosen topic with the topic assigned to each of the VideoGameGeek games the user has played in the past.

After an LDA has been trained for the topic analysis of our recommendation system, we can determine which topic is to be assigned to the game by using Equation 4.4, which is

arcade 0.0168, version 0.0164, fighter 0.0153, super 0.0124, street 0.0117, characters 0.0110, fighting 0.0075, time 0.0069, turtles 0.0061
Teenage Mutant Ninja Turtles: Turtles in Time is a video game produced by Konami. A sequel to the original Teenage Mutant Ninja Turtles (TMNT) arcade game, it is a scrolling beat 'em up based on the 1987 TMNT animated series. Originally an arcade game, Turtles in Time was ported to the Super Nintendo Entertainment System (SNES) as Teenage Mutant Ninja Turtles IV: Turtles in Time in 1992. That same year, a similar game, Teenage Mutant Ninja Turtles: The Hyperstone Heist was released for Sega Mega Drive/Genesis. Years later, the arcade version of Turtles in Time was revisited on newer consoles. A slightly altered version of the arcade game was included as an unlockable bonus in the 2005 game Teenage Mutant Ninja Turtles 3: Mutant Nightmare. In August 2009, Ubisoft released a 3D remake of the game, Teenage Mutant Ninja Turtles: Turtles in Time Re-Shelled, for Xbox Live Arcade. The remake was released onto PlayStation Network on September 10, 2009. Players guide the turtles through a series of levels, starting out in the streets of New York City before being transported to levels representing various eras of history.

Figure 4.6: Topic assigned to the game "Teenage Mutant Ninja Turtles: Turtles in Time" and its description with the matching keywords highlighted

given as an input a game description D. Based on the distribution of each word w_i in D and the probability of occurrence of each word w_i in each topic z_j, i.e., $P(w_i|z_j)$, Equation 4.4 determines the topic with the*highest* probability for the game description.

$$Topic(D_G) = max_{j=1}^{N} \sum_{i=1}^{K} P(w_i|z_j) \tag{4.4}$$

where $Topic(D_G)$ is the topic assigned to the description D of game G, N is the total number of latent topics, i.e., 27 in our case, K is the total number of words in D_G, $P(w_i|z_j)$ is the probability of word w_i given the topic z_j.

An example for the topic assigned to the VidoGameGeek game "Teenage Mutant Ninja Turtles: Turtles in Time" along with its description is shown in Figure 4.6.

4.3.3 Predicted Ratings of Therapeutic Games

VideoGameGeek posts *user ratings* on games and the average user rating of each game. Based on these ratings, we can predict the user rating for each therapeutic game. The predicted rating is then used to partially rank the therapeutic games to be played by autistic adults.

After obtaining the *average rating* of the 10 VideoGameGeek games provided by an autistic adult, we apply our *rating prediction algorithm* to rate therapeutic games to be recommended to the adults based on those ratings. Our rating prediction algorithm relies on a weighing mechanism to decide the *closeness* between a therapeutic game and each of the 10 VideoGameGeek games based on the genres of the games. We apply a metric called *word correlation factor* (defined in Section 4.3.3) to *weigh* the games using *genres* and then rank each therapeutic game with respect to the 10 VideoGameGeek games.

Rating of VideoGameGeek

VideoGameGeek includes over 45,000 games and each of those games is associated with an *average user rating*. Players of a VideoGameGeek game rate the game from 1 to 10, with 10 being the highest, and an average is computed to obtain a single rating for the game. The following table shows the average rating of some VideoGameGeek games:

Game	**Average VideoGameGeek Rating**
Minecraft	7.93
Angry Birds	6.55
Grand Theft Auto	6.61

Word Correlation Factor (WCF)

Using the average ratings of VideoGameGeek games and determining how closely related a VideoGameGeek game and a therapeutic game are based on the *word-correlation factors* of the genres of the games, we can predict the rating of the therapeutic game. *Word correlation factor (WCF)* is a measure to capture the closeness of two *non-stop, stemmed words*. WCF is computed using the *frequency of co-occurrence* of two non-stop, stemmed words and the *distance* between the two words in a document. The distance between two words are the number of words between the two words. To calculate the correlation factor between two non-stop, stemmed words based on a collection of 880,000 Wikipedia documents, we apply the Equation 4.5.

$$WCF(i,j) = \frac{\sum_{W_i \in V_i} \sum_{W_j \in V_j} \frac{1}{d(W_i, W_j)+1}}{|V_i| \times |V_j|} \tag{4.5}$$

where $d(W_i, W_j)$ is the *distance* between two words W_i and W_j in a Wikipedia document, V_i and V_j are the stem[1] variations of the words i and j, respectively, in the document. $|V_i|$ refers to the number of stem variations of word i and $|V_j|$ refers to the number of stem variations of word j.

WCF is used to determine the *weight* between a therapeutic game and a VideoGameGeek game based on the *genre* of the games. These weights are used in combination with the *ratings* (and other attributes values to be discussed in the subsequent sections) of those games to assign ratings for the therapeutic games. Since there can be multiple keywords associated with a genre, we combine the weights for each of pair of keywords extracted from two genres to create the cumulative weight using Equation 4.6.

$$Cumulative\ Weigth(TG_p, VG_q) = \frac{\sum_{i=1}^{n} Min(\sum_{j=1}^{m} WCF(TG_{p_i}, VG_{q_j}), 1)}{n} \tag{4.6}$$

where *cumulative weight*(TG_p, VG_q) is the *weight* for the p^{th} genre of therapeutic game TG and the q^{th} genre of VideoGameGeek game VG, $WCF(TG_{p_i}, VG_{q_j}$ is the *word correlation factor* between the i^{th} keyword in the p^{th} genre of TG and the j^{th} keyword in q^{th} genre of VG, n is the total number of keywords in the p^{th} genre, m is the total number of keywords in the q^{th} genre. The Min function is introduced to constrain the value of the WCF within 0 and 1. In case of *exact match* of keywords, their WCF value is 1 which makes the *cumulative weight* of a genre greater than 1. In order to avoid domination of one keyword over other, the Min function makes sure that the weight is between 0 and 1.

Since a single game can have multiple genres and each genre of a particular game has a single *cumulative weight* with respect to another game, we obtain a single value using these cumulative weights for each game and call it *actual weight*, as it is the *weight* which is

[1] Stemmed word refers to a word which has been changed into its grammatical root. For example, 'assign' is the stem of words 'assignment', 'assigned', 'assigning', etc.

4X Strategy, Action, Action Adventure, Action RPG, Adventure, Arcade, Augmented Reality, Beat 'em up, Classic Games, Clicker, Dating sim, Dungeon Crawler, Educational, Endless runner, Fighting, First person shooter, Fitness, Flight simulator, Hidden object, Interactive movie, Life simulation, Light gun shooter, Mana gement, Maze, MMO, MOBA, Other, Party, Pinball, Platform, Point-and-click, Puzzle, Racing, Real time strategy, Rhythm, Rouguelike, RPG, Run-n-gun, Sandbox, Scrolling, Shoot 'em up, Shooter, Simulation, Sports, Stealth, Strategy, Text adventure, Tower defense, Trivia, Visual novel, Walking simulator

Figure 4.7: Genres in VideoGameGeek

actually used in partially determining the *rating* of a therapeutic game. The actual weight of a VideoGameGeek game with respect to a therapeutic game is defined by Equation 4.7.

$$Actual\ Weight(TG, VG) = \frac{\sum_i^P \sum_j^Q Cumulative\ Weight(TG_i, VG_j)}{|P| \times |Q|} \tag{4.7}$$

where $Actual\ Weight(TG, VG)$ is the actual weight of VideoGameGeek game VG with respect to therapeutic game TG, $Cumulative\ Weight(TG_i, VG_j)$ is the cumulative weight of i^{th} genre of TG with respect to the j^{th} genre of VG as defined in Equation 4.6, $|P|$ is the total number of genres in TG and $|Q|$ is the total number of genres in VG.

The Genre-Based Rating Prediction approach

After we have obtained *actual weight* of the VideoGameGeek games corresponding to a therapeutic game, we use the collaborative filtering technique to *rate* the therapeutic game based on the average of rating of VideoGameGeek games along with the actual weights of VideoGameGeek games with respect to the therapeutic game. VideoGameGeek comes with 53 *unique* genres. These genres describe the nature of the games and are shown in Figure 4.7.

Given the top-10 VideoGameGeek games a user likes to play, we obtain the genres for each of these games from VideoGameGeek. We also assign genres to the therapeutic games. The weight of each VideoGameGeek game and a therapeutic game is determined using *actual weight*, which indicates how closely related the two games are. We obtain the predicted therapeutic *game rating* using (i) the weights between the therapeutic game and

Input: 10 VideoGameGeek games provided by the user called VideoGameGeek Games, and a therapeutic game T
Output: The predicted rating for the therapeutic game

1. *Initialize W, as the size of 1X10, and R as 1X10 as there are 10 VideoGameGeek games*
2. *For each game VG in VideoGameGeek Games*
    - i. *Assign W[VG] as the actual weight between T and VG obtained by using WCF between the keywords of the genres in T and VG and then using cumulative weight between each genre*
    - ii. *Assign R[VG] as the rating of VG*

  b. *Calculate rating for T using W and R using Equation 4*
  c. *Return rating for T*

Figure 4.8: Algorithm to predict rating for a therapeutic game

the 10 VideoGameGeek specified in a user profile based on their genres, and (ii) the average rating for each VideoGameGeek game using Equation 4.8.

$$r_{u,k} = \frac{\sum_{j \in N_u} W_{j,k} \times r_j}{\sum_{j \in N_u} W_{j,k}} \tag{4.8}$$

where $r_{u,k}$ is the *predicted rating* of the therapeutic game k for user u, $w_{j,k}$ is the *actual weight* of the therapeutic game k and one of the VideoGameGeek games j specified in u's profile based on their genres, r_j is the *average rating* for the VideoGameGeek game j, N_u is the number of VideoGameGeek games specified in u's profile, which is 10.

After the *rating* for each of the therapeutic games is predicted for a particular user, we can *partially rank* those games to be recommended to the user. The algorithm for genre-based rating prediction is shown in Figure 4.8.

To illustrate the rating prediction, we assume a user U who likes to play three VideoGameGeek games, namely 'Minecraft', 'Angry Birds', and 'Grand Theft Auto'. The ratings of the games are obtained from VideoGameGeek and are shown in Table 4.3. The genre of 'Minecraft' is *Sandbox* and *Survival*, 'Angry Birds' is *Puzzle*, and that of 'Grand Theft Auto' is *Action Adventure*. Further assume that the therapeutic game for which we are going to predict the ranking is 'Zombie Maze'. The genre of 'Zombie Maze' is *Action Adventure* and *Survival*.

The genres of the three VideoGameGeek games along with their *cumulative* and actual weights with respect to the genres of 'Zombie Maze' is shown in Table 4.3. After the WCF scores of each genre keyword have been calculated, the scores can be used to calculate the *cumulative weight* of each genre based on Equation 4.6, followed by the *actual weight* for each game using Equation 4.7.

The final rating and weights of the VideoGameGeek games are shown in Table 4.4. Given the weights and the ratings, we apply Equation 4.8 (genre-based rating prediction) to predict the *rating* for 'Zombie Maze', which is **7.23**. Using this approach, the *predicted rating* for other therapeutic games for user U are calculated.

Number of Players

The number of players of a game is a range of potential players who can play the game. Some of the VideoGameGeek and therapeutic games involve multiple players while others are just single-player games. We compare the recommended number of players in each VideoGameGeek game specified in a user profile with that of a therapeutic game to rank the latter using Equation 4.9.

$$NPlayers(TG) = \begin{cases} \sum_{n_{TG}} P_{TG} + \sum_{n_{VG}} -1 \times P_{VG} & \text{if } n_{TG} = 1 \\ -(\sum_{n_{TG}} P_{TG} + \sum_{n_{VG}} -1 \times P_{VG}) & \text{if } n_{TG} > 1 \end{cases} \tag{4.9}$$

where $NPlayers(TG)$ is the score based on the potential number of players for the therapeutic game TG, n_{TG} is the number of player options in TG, n_{VG} is the number of player options in the VideoGameGeek game VG. P_{TG} is (one of) the possible number(s) of players in TG, and P_{VG} is (one of) the possible recommended number(s) of players in VG.

Based on the score obtained for each of the therapeutic games as computed by using Equation 4.9, we rank the therapeutic games in *decreasing order* such that the therapeutic game that receives the highest score with respect to the VideoGameGeek games in a user profile is ranked as the highest.

Final ranking of therapeutic games

After we have obtained the partial ranking for each therapeutic game based on one of its four different attribute values, we combine them to obtain a single ranking for each of the therapeutic games. The single ranking of a therapeutic game determines the cumulative effect of all the four attributes, and we rely on the CombMNZ model to obtain the final ranking of the game. CombMNZ is a commonly-used fusion method [3] that combines multiple ranking lists on an item I to determine a *joint ranking* of I and is calculated as

$$CombMNZ_I = \sum_{K=1}^{N} I^K \times |I^K > 0| \tag{4.10}$$

where N is the number of ranked lists to be fused, which is *four* in our case, I^K is the normalized score of item I in the ranked list K, and $|I^K > 0|$ is the number of non-zero, normalized scores of I in the ranked list K.

Prior to using Equation 4.10, we normalize the original score in each attribute rank list of a therapeutic game into a common range $[0, 1]$ using Equation 4.11.

$$I^K = \frac{S^I - I^K_{min}}{I^K_{max} - I^K_{min}} \tag{4.11}$$

where S^I is the score of I in the ranked list K to be normalized, I^K_{max} (I^K_{min}, respectively) is the maximum (minimum, respectively) score available in K, and I^K is the normalized score for I in K.

Table 4.1: Correlation of questions in questionnaire with weakness areas AC (Developing Audio Communication Skills), EC (Maintaining Eye Contact), FE (Recognizing Facial Expression), SE (Showing Empathy), and ST (Engaging in Speech Therapy)

Questions	**Weakness Area**
When I talk on the phone, I'm not sure when it's my turn to speak	AC, ST
I am often the last to understand the point of a joke	ST, SE
I often offend people unintentionally	SE
Other people frequently tell me that what I've said is impolite, even though I think it is polite	SE
I get upset when my work is criticized	SE
I become overwhelmed when looking into someone's eyes	EC
I know how to tell if someone listening to me is getting bored	FE
When I talk, it isn't always easy for others to get a word in edgewise	ST
I am quick to comfort someone in distress	SE
I enjoy getting feedback	ST, SE
People describe me as clumsy	ST
I often notice small sounds when others do not	ST
I frequently get into arguments	SE
I find it hard to tell if someone is interested	FE
When I'm reading a story, I can easily imagine what the characters might look like	FE
I look around the room while I talk	EC
I am able to resolve conflicts easily	SE
My friendships never last long	SE, AC, FE, ST
I find social situations easy	SE, AC, FE, ST, EC
I find it easy to work out what someone is thinking or feeling just by looking at their face	FE
In a social group, I can easily keep track of several different people's conversations	AC
I enjoy meeting new people	AC, FE, ST, EC
I frequently find that I don't know how to keep a conversation going	AC
I am good at social chitchat	AC
I find it hard to make new friends	AC
I am a good diplomat	AC, FE, EC
I find it difficult to work out people's intentions	FE, EC
I prefer to do things with others rather than on my own	AC
I get upset from disagreements with my friends	SE
I don't know what to say when introducing myself	AC
I avoid eye contact when talking to people	EC
I know all the social rules	AC, EC, FE
I find it easy talk and make eye contact at the same time	EC
I find it easy to 'read between the lines' when someone is talking to me	EC

Table 4.2: VideoGameGeek game attributes

Game Attribute	**Description**	**Example**
Theme	The category to which the game belongs	Cartoon, Comedy
Topic Analysis	The topic associated to the game	Star Wars
Platform	The application in which the game is played	Windows, Android
No. of players	Maximum no. of players who can play	2

Table 4.3: Genres of example games with cumulative and actual weights, where *Cum* stands for *cumulative*

Zombie Maze	**Minecraft**			**Angry Birds**			**Grand Theft Auto**		
Genre	*Genre*	*Cum Weight*	*Actual Weight*	*Genre*	*Cum Weight*	*Actual Weight*	*Genre*	*Cum Weight*	*Actual Weight*
Action Adventure	Sandbox	0.3	0.52	Puzzle	0.2	0.25	Action Adventure	1.0	0.8
	Survival	0.4							
Survival	Sandbox	0.4		Puzzle	0.3		Action Adventure	0.6	
	Survival	1.0							

Table 4.4: Predicted rating for the game 'Zombie Maze'

Metrics	**Minecraft**	**Angry Birds**	**Grand Theft Auto**
Weight	1.0	0.25	0.8
Rating	7.93	6.55	6.61

Chapter 5

Recommendation of Games Using Accumulative Baseline Values

After sequencing therapeutic games using VideoGameGeek attribute values, we ask the users to play those games, starting with the top-ranked game, and continue until certain criterion is satisfied in each weakness area prevalent in them.

5.1 Baseline Scores

After the user plays a game belonged to a particular weakness area, (s)he is asked to answer the questionnaire again and the *initial baseline score* of that weakness area is modified to become the *updated baseline score* for that area. The updated baseline score shows the impact of playing the game with respect to the user. If the updated baseline score for an area *increases* compared with the initial baseline score, we claim that the corresponding game has *affinity* to that user. If the updated baseline score for that area is at least 5.5, which is the threshold, we assume that the user has made significant improvement in that weakness area. We set the threshold as 5.5 because the scores to the answers range from 1 to 7 and 5.5 lies slightly towards the right of the average, indicating the complete absence of a weakness area. After the user crosses the threshold, the weakness area is *removed* from the list of weakness areas designated for the user. If the updated baseline score does not exceed 5.5, we cannot be certain that the weakness area has been fully addressed. Therefore, we record the original value of the initial baseline score, called *original initial baseline score*, as it will be used later and assign the initial baseline score of that weakness area as updated baseline score. Hereafter, we recommend another game from the ranked list of games in that weakness area to the user. We continue this process for a user until either the threshold of 5.5 for each weakness area is reached or we exhaust all the games in each weakness area.

5.2 Accumulative Values

While the users are playing therapeutic games recommended to them, we maintain a key metric, called *accumulative values*, which captures the change in the baseline scores of each particular weakness area after the user plays the therapeutic games in the corresponding area. The accumulative values of a user is a vector of size $1 \times N$, where N is the total number of therapeutic games and each component of the vector stores an accumulative value corresponding to a therapeutic game (in a particular weakness area). The index of the values for games and the ordering of the therapeutic games in each weakness areas in the accumulative values remains the same for all the users. The layout of therapeutic games in each weakness area is shown in below. After the user plays a therapeutic game, the *accumulative value* for that game is set as the *difference* between the *updated baseline score* and the *initial baseline score* of the weakness area. For the therapeutic games that are *never* played or that do *not* show any change between the two scores, the accumulative value for the game remains *zero*.

Developing Audio Communication Skills	**Maintaining Eye Contact**	**Recognizing Facial Expression**	**Showing Empathy**	**Engaging in Speech Therapy**
$G_{1_1}, \ldots, G_{1_{M_1}}$	$G_{2_1}, \ldots, G_{2_{M_2}}$	$G_{3_1}, \ldots, G_{3_{M_3}}$	$G_{4_1}, \ldots, G_{4_{M_4}}$	$G_{5_1}, \ldots, G_{5_{M_5}}$

5.3 Affinity Values

After all the weakness areas prevalent in a user have been addressed, i.e., either the user has played each therapeutic game in the weakness area or (s)he has made significant improvement (crossed the threshold) in that area, we assign **affinity values** for the user, which is the *latest accumulative values* of all the therapeutic games. After the *affinity values* is generated for a user profile, the user profile is called an *updated profile*, and the updated profile is used

```
Algorithm: Calculating the affinity values
Input: The initial baseline scores of a user U represented as initial_baseline_score of size 1X5, list L of weakness areas of the user, the list of therapeutic games, questionnaire with K questions whose answers range from 1 to 7
Output: A vector, the affinity_values for user U
  1. Initialize the updated_baseline_scores as [0,0,0,0,0] of size 1X5, and accumulative _values of size 1XN as
     [0,0,0,0,0….0], where N is the total number of therapeutic games
  2. For each weakness area W in L
     2.1. Initialize threshold_tracker as 0
     2.2. While (threshold_tracker < 5.5 or there is no game remaining in W)
          2.2.1. Ask the user to play G in W based on its ranking in W for U
          2.2.2. Ask the user to fill up the questionnaire
          2.2.3. Initialize current_score as 0
          2.2.4. For each answer A belonged to the questions in W in the questionnaire
                     current_score := current_score + score for A
          2.2.5. updated_baseline_scores[W] := current_score / number of questions in W
          2.2.6. accumulative _values[G] := updated _baseline_scores[W] -
                 initial_baseline_scores[W]
          2.2.7. initial_baseline_scores[W] := updated_baseline_scores[W]
          2.2.8. threshold_tracker := updated_baseline_scores [W]
          2.2.9. Remove G from W
  3. Return accumulative _values as affinity_ values
```

Figure 5.1: Algorithm 2. Calculation of affinity values

in the 2^{nd} phase for recommendation. The layout of updated profile for a user is shown in below, where $BS_1, \ldots, BS_5$ represent the *original initial baseline scores* in all the weakness areas, $G_1A_1, \ldots, G_{10}A_4$ is a matrix of size 10×4 representing the *attribute values* of each of the top-10 VideoGameGeek games preferred by the user, and $AS_1, \ldots, AS_n$ are the *affinity values*, where n is the total number of therapeutic games.

Original Baseline Scores for the Weakness Areas					Attribute Values Obtained from VideoGameGeek				Affinity Values			
BS_1	BS_2	BS_3	BS_4	BS_5	G_1A_1	G_1A_2	G_1A_3	G_1A_4	AS_1	AS_2	$\ldots$	AS_n
					G_2A_1	G_2A_2	G_2A_3	G_2A_4				
					$\ldots$	$\ldots$	$\ldots$	$\ldots$				
					$G_{10}A_1$	$G_{10}A_2$	$G_{10}A_3$	$G_{10}A_4$				

Figure 5.1 shows the algorithm that generates the *affinity values* for a user and an example of the algorithm is shown below. Based on the vectors of affinity values for a number of users, we can determine which games can be effectively recommended to new users during the 2^{nd} phase of our recommendation process.

Example 1 Assume that Alice has three weakness areas { 'Developing Audio Communication Skills', 'Recognizing Facial Expression', and 'Maintaining Eye Contact' }. Let the *initial baseline scores* of Alice before playing any therapeutic game be [1, 3, 2], and altogether there are five therapeutic games.

Let the ranked list of therapeutic games to be considered for recommendation to Alice be

- Developing Audio Communication Skills $\rightarrow$ { Communication Quiz, Zombie Maze }
- Recognizing Facial Expression $\rightarrow$ { Jim's World, Make New Friends }
- Showing Empathy $\rightarrow$ { Voice }

Accumulative values for Alice is initially set to {0, 0, 0, 0, 0} for {Communication Quiz, Zombie Maze, Jim's World, Make New Friends, Voice}.

To begin with, Alice plays 'Communication Quiz' and answers the questionnaire again and assume that the *new baseline score* for 'Developing Audio Communication Skills' is 6. Since the current baseline score has increased to greater than 5.5, 'Communication Quiz' has *positive affinity* to Alice. We assign the difference, which is 5, between *new* and *initial baseline score* as the *accumulative value* of 'Communication Quiz'. Therefore, Alice's *Accumulative_values* is {5, 0, 0, 0, 0}.

Alice is no longer required to play the game 'Zombie Maze', since she has already crossed the threshold for 'Developing Audio Communication Skills'. Alice then plays 'Jim's World' and assume that the *new baseline score* for the 'Recognizing Facial Expression' weakness area increases to 4. Therefore, Alice's updated *Accumulative_values* is {5, 0, 1, 0, 0}, since the value for 'Recognizing Facial Expression before playing 'Jim's World' was 3 and now that the baseline score has increased to 4. We store the difference, i.e. 1, in the 3^{rd} component of the *accumulative values* as 'Jim's World' is the 3^{rd} index.

As the *current baseline score* of the 'Recognizing Facial Expression' has not crossed over the threshold, Alice then plays 'Make New Friends' and assume that her *new baseline* *score* in that area becomes 7, which is greater than 5.5. Hence, Alice's *updated Accumulative_values* is {5, 0, 1, 3, 0}, since the difference between old baseline score, i.e., 4, and the new baseline score, i.e., 7, is 3.

Further assume that Alice plays 'Voice' and her new baseline score for the Showing Empathy area increases to 6. As a result, Alice's *updated Accumulative_values* is {5, 0, 1, 3, 4}.

Therefore, Alice's *affinity_values* is {5, 0, 1, 3, 4}. □

Chapter 6

An Automated Game Recommendation Strategy

In this, the 2^{nd} phase, of our recommendation process, we develop an automated recommending mechanism in which, given the profile of each new user, we use the data from updated profiles of former users to rank the list of therapeutic games to be recommended to the user. By assigning affinity values to users in Phase 1, we collected the required data for this automated recommendation approach. This recommendation strategy is fully automated because we do not require users to fill up the questionnaire again and again, but only for the first time, to determine the therapeutic games recommended for the user to play. In order to automate the process of recommendation, we determine the *affinity values* for a new user by itself so that the user can be recommended games automatically based on the strategy discussed in Section 6.2. We adopt KNN (K-Nearest Neighbor) regression for calculating affinity values for the new user.

6.1 KNN Regression

KNN regression is a supervised learning algorithm [13] which is used to find the output label or class for a new data point based on its closest neighbors. Supervised learning algorithms are divided into classification and regression [7] and we adopt regression because we are supposed to estimate the value of output label for each new user by calculating the mean of output labels of its neighboring data points. (The output label for each data point is further discussed in Section 6.2.) For each new user U, KNN regression calculates the average of the output label of the K neighbors who are closest to U and assign as the class to U. The K in this approach refers to the numbers of neighbors that is used to classify the new data point into a category or a class and is determined empirically. We do not consider other clustering algorithms [17] for calculating affinity values for a new user because they are mostly unsupervised, i.e., they do not have an output label. Since we already have labeled classes, i.e., the affinity values of different users, we adapt KNN, a supervised learning approach. KNN is preferred over other supervised learning approaches which include multilayer perceptron, Support Vector Machine (SVM), linear regression, etc., because KNN is an instance-based learning algorithm which means it adapts to new addition of data and does not require training to build any model.

6.2 Calculating Affinity Values with KNN

As KNN regression assigns output label to a data point by averaging the labels of its neighbors, which are the *affinity values* of the updated profiles that are closest to the new user profile in our approach, the affinity values for a new user could be computed by averaging the affinity values of those updated profiles. In order to train a KNN regression model to obtain the average of affinity values to be assigned as the affinity values for the new user, we consider *features* which include the *initial baseline scores* of a new user and the *attribute values* of the top-10 VideoGameGeek games that the user prefers. We use these attribute values because it is likely that similar players play games that are assigned same VideoGameGeek game attribute values. The data for each feature that is used in KNN is shown in Table 6.1.

The motivation behind using these *features* for running KNN is that if the initial baseline scores of two users are close, and if they have played same types of games in the past, they could be facing challenges in the same weakness areas. For example, if players prefer to play single-player game only, it could mean that they are less interested in playing with others. We change the representation of the features to numerical and normalize them so that we can calculate *distance* between them as vectors in the vector space. We find the distance between feature values of updated profiles with that of the new user profile and sort

Table 6.1: Features of KNN used in our automated game recommendation approach

Features	**Source**	**Examples**
Baseline Score for *Developing Audio Communication Skills*	Questionnaire Answered	3
Baseline Score for *Maintaining Eye Contact*	Questionnaire Answered	2
Baseline Score for *Recognizing Facial Expression*	Questionnaire Answered	1
Baseline Score for *Showing Empathy*	Questionnaire Answered	5
Baseline Score for *Engaging in Speech Therapy*	Questionnaire Answered	6
Themes of Games	VideoGameGeek	Fighting
Number of players	VideoGameGeek	3
Topic analysis of Games	VideoGameGeek	Animals
Platforms of Games	VideoGameGeek	Windows

them in increasing order using the *Euclidean distance* measure. We then take the *average* of affinity values for the first K updated profiles and assign it as the *affinity values* for the new user because the label for our KNN regression model is an affinity values vector. Figure 6.1 shows Algorithm 3, the KNN algorithm for the automated recommendation generation.

6.3 Using Affinity Values to Recommend Games

When a new user U fills up the questionnaire, we can determine which weakness areas are prevalent in the user according to the *initial baseline scores*. The user is recommended therapeutic games in those areas based on the decreasing order of affinity value of the games assigned to U using Algorithm 2 in each area. Shown below is the layout of the affinity values obtained for a user by running KNN, where G_{i_j} is the therapeutic game in each weakness area, V_{i_j} is the affinity value for each game in a weakness area, $1 \leq V_{i_j} \leq 7$, $1 \leq i \leq 5$, $1 \leq j \leq x_k$, and x_k represents the number of therapeutic games in k^{th} $(1 \leq k \leq 5)$ weakness area.

Weakness Area 1			Weakness Area 2			Weakness Area 3			Weakness Area 4			Weakness Area 5		
G_{1_1}	$\ldots$	$G_{1_{x_1}}$	G_{2_1}	$\ldots$	$G_{2_{x_2}}$	G_{3_1}	$\ldots$	$G_{3_{x_3}}$	G_{4_1}	$\ldots$	$G_{4_{x_4}}$	G_{5_1}	$\ldots$	$G_{5_{x_5}}$
V_{1_1}	$\ldots$	$V_{1_{x_1}}$	V_{2_1}	$\ldots$	$V_{2_{x_2}}$	V_{3_1}	$\ldots$	$V_{3_{x_3}}$	V_{4_1}	$\ldots$	$V_{4_{x_4}}$	V_{5_1}	$\ldots$	$V_{5_{x_5}}$

There can be a discrepancy between the weakness areas shown by the *initial baseline scores* and the *affinity values* of the user, and there are many different scenarios to consider

Algorithm: K-Nearest Neighbor for affinity values generation
Input: Number of updated profiles M, Initial baseline scores for each M updated profiles IBS_UP, attribute values for the top-10 games for each M updated profiles represented as AT_UP, whose size is MX10X4 since there are 4 attribute values for each of the top-10 game, initial baseline scores for a new user IBS_U, attribute values of the top-10 games of new user AT_U
Output: The Affinity values for new user U

1. *Initialize an array L of size 1XM, Baseline_distance with size 1xM as 0, and Affinity_value_distance with size 1xM as 0*
2. *Initialize counter as 1*
3. *While (counter <= M)*
    a. *Baseline_distance[counter] := Euclidean distance between IBS_UP[counter] and IBS_U*
    b. *Increment counter by 1*
4. *Re-initialize counter as 0*
5. *While (counter <= M)*
    a. *Initialize game_counter as 1*
    b. *For (game_counter := 1 to 10)*
        i. *Initialize new_user_game_counter as 0*
        ii. *For (new_user_game_counter := 1 to 10)*
            1. *Affinity_value_distance[counter] := Affinity_value_distance[counter] + Euclidean distance between AT_UP[game_counter] and AT_U[new_user_game_counter]*
            2. *Increment new_user_game_counter by 1*
    c. *L[counter] := Baseline_distance[counter] + Affinity_value_distance[counter]*
    d. *Increment counter by 1*
6. *Sort L in increasing order*
7. *Take top K updated profiles as the closest neighbors of U*
8. *Take the mean of affinity values of the K-nearest neighbors, denoted as AS*
9. *Return AS*

Figure 6.1: Algorithm 3: K-Nearest Neighboring Algorithm for affinity values calculation of new users

while recommending games to the user due to this discrepancy which are listed as cases below.

- Case 1. If the weakness areas shown by both initial baseline scores and the affinity values are the *same*, we simply rely on the affinity value of each game in those weakness areas to recommend games based on the *descending order* of the *affinity values*. However, sometimes two or more therapeutic games in the same weakness area may have the same affinity value. In this case, we break the *tie* to rank those games using the *ranking* of the games as done in Phase 1.

- Case 2. The initial baseline scores can also indicate the presence of *larger* number of weakness areas than the affinity values. For the weaknesses not shown by affinity values, we rely on the ranking of therapeutic games generated by comparing them with the VideoGameGeek games the user likes to play.
- Case 3. The affinity values can indicate *larger* number of weakness areas than that shown by the initial baseline scores. In this case, we do not recommend games in the weakness areas that are not indicated by the initial baseline scores.

To illustrate different cases, the set of weakness areas prevalent in the user as shown by the *initial baseline scores* is represented as T and the set of weaknesses area prevalent in the user as shown by his/her *affinity values* is represented as S. If the *initial baseline scores* and *affinity values* indicate same weakness areas, these areas are included as $T \cap S$. Those weakness indicated only by *initial baseline scores* is given by $T - (T \cap S)$ and those weakness areas indicated only by the *affinity values* is given by $S - (T \cap S)$.

6.3.1 Deficiencies in $T \cap S$

For the deficiencies in $T \cap S$, i.e., the weakness areas shown as existent in the user by both his/her *initial baseline scores* and *affinity values*, we recommend games based on the decreasing *affinity values* of therapeutic games in each weakness area.

Let us consider an example where there are 5 weakness areas and there are two therapeutic games in each of those areas as shown below. Assume that the set of weaknesses indicated by *initial baseline scores*, T, is {'maintaining eye contact', 'showing empathy', 'engaging in speech therapy'}. Further assume that the set of weakness areas indicated by the *affinity values*, S, is also {'maintaining eye contact', 'showing empathy', 'engaging in speech therapy'}. Hence, we recommend therapeutic games to the user in the three weakness areas, i.e., 'maintaining eye contact', 'showing empathy', and 'engaging in speech therapy'. We consider the affinity value for each game in the same area to rank the games. As shown below, the value for G_{2_2} is greater than G_{2_1} for 'maintaining eye contact'. Therefore, G_{2_2}

is recommended first. Similarly, since G_{4_2} has higher value than G_{4_1} for 'showing empathy', it is recommended first, and G_{5_2} is recommended, since it is ranked higher than G_{5_1} for 'engaging in speech therapy'.

Developing Audio Communication Skills		Maintaining Eye Contact		Recognizing Facial Expression		Showing Empathy		Engaging in Speech Therapy	
G_{1_1}	G_{1_2}	G_{2_1}	G_{2_2}	G_{3_1}	G_{3_2}	G_{4_1}	G_{4_2}	G_{5_1}	G_{5_2}
0	0	2	3	0	0	1	5	1	6

Tie Breaker

Multiple therapeutic games in the same weakness area can have the same affinity value. Since we recommend games to the user based on the *descending order* of the *affinity values*, there needs to be a *tie-breaking* mechanism to generate a sequence of game to be recommended to the user. In case of a *tie* in the affinity values, we consider the ranking of the games obtained by using the ranking mechanism in Phase 1, i.e., based on the *attribute values* of VideoGameGeek. The game ranked higher by the user is given a higher priority.

Considering the following example when the affinity values for therapeutic games are the same in 'Showing empathy'. Therefore, we recommend games in the area by *ranking* the therapeutic games by comparing them with the VideoGameGeek games played by the user in the past.

Developing Audio Communication Skills		Maintaining Eye Contact		Recognizing Facial Expression		Showing Empathy		Engaging in Speech Therapy	
G_{1_1}	G_{1_2}	G_{2_1}	G_{2_2}	G_{3_1}	G_{3_2}	G_{4_1}	G_{4_2}	G_{5_1}	G_{5_2}
0	0	2	3	0	0	4	4	2	3

6.3.2 Deficiencies in $T - (T \cap S)$

When the user fills up the questionnaire for the first time, his/her *initial baseline scores* may indicate that they have weakness in areas that are not indicated by the *affinity values*. These areas of weakness are included under $T - (T \cap S)$. The situation may arise when the new user actually has deficiencies in those areas, but his neighbors do not. In such case, we adopt the methodology in Phase 1 to recommend games to the user.

Let us consider the following set of weakness areas indicated by *initial baseline scores* for a user, where T is {'Maintaining Eye Contact', 'Recognizing Facial Expression', 'Showing Empathy', 'Engaging in Speech Therapy'}. Let the set of weakness area indicated by his *affinity values*, S, be {'Maintaining Eye Contact', 'Showing Empathy'}. In this situation, we adopt the approach in Phase 1 to recommend games in 'Recognizing Facial Expression‘, and 'Engaging in Speech Therapy'.

Developing Audio Communication Skills		Maintaining Eye Contact		Recognizing Facial Expression		Showing Empathy		Engaging in Speech Therapy	
G_{1_1}	G_{1_2}	G_{2_1}	G_{2_2}	G_{3_1}	G_{3_2}	G_{4_1}	G_{4_2}	G_{5_1}	G_{5_2}
0	0	2	3	0	0	4	6	0	0

6.3.3 Deficiencies in $S - (T \cap S)$

The affinity values of a new user may show some weakness areas prevalent in the user while his/her initial baseline scores may not show them. The situation arises when the new user himself does not have the corresponding weakness, but some of his nearest neighbors do. For these weakness areas, we do not ask the new user to play any games as the user has already indicated that (s)he does not have that weakness in the *initial baseline scores*.

For example, assume that the *affinity values* for that user appears as shown below, where S = {'Developing Audio Communication Skills', 'Maintaining Eye Contact', 'Recognizing Facial Expression', 'Showing Empathy'}, and the *initial baseline scores* shows that T = {'Maintaining Eye Contact', 'Recognizing Facial Expression'}. We do not recommend the user to play games in the weakness areas included in $S - (T \cap S)$, i.e., 'Developing Audio Communication Skills' and 'Showing Empathy'.

Developing Audio Communication Skills		Maintaining Eye Contact		Recognizing Facial Expression		Showing Empathy		Engaging in Speech Therapy	
G_{1_1}	G_{1_2}	G_{2_1}	G_{2_2}	G_{3_1}	G_{3_2}	G_{4_1}	G_{4_2}	G_{5_1}	G_{5_2}
1	6	2	3	2	5	4	6	0	0

Chapter 7

Experimental Results

In this chapter, we present the performance evaluation of our gaming and recommendation system. We first discuss the experiments conducted during the 3-month period, which is followed by the data used for the empirical study. Hereafter, we introduce the statistical method employed to verify the usefulness of our gaming and recommendation system known as the Wilcoxon test, which is followed by the performance analysis of the empirical study conducted on our gaming and recommendation system and the examination on the statistical results based on the study.

7.1 The Empirical Study

In order to perform the evaluation to verify the novelty and merit of our gaming and recommendation system, we conducted an empirical study to determine the impact of playing recommended therapeutic games by adults with ASD. The study included (i) arranging different groups of adults with ASD who participated in performing different tasks, (ii) recommending therapeutic games to those who played games based on their areas of weakness, and (iii) observing the skill changes in the adults after playing the games. Since the ultimate goal of our gaming and recommendation system is to suggest games to autistic adults so that they can improve their social-interactive skills, we can draw a conclusion on whether our goal has been met based on the empirical study.

We collaborate with ScenicView Academy for conducting our empirical study. ScenicView is a non-profit school for adults with ASD and learning disabilities located in Provo, Utah. Their program is "centered on empowering students with skills to live an independent, productive, fulfilling life" [ScenicView]. The school was established in 2000 and provides residential and day programming to young adults diagnosed with ASD, Executive Functioning Deficits, Dyslexia, Dyscalculia, Dysgraphia and other learning differences. A mission of ScenicView is to improve daily living skills of the adults like cooking, transportation, cleaning, etc., and lifelong educational skills such as writing and learning. Since our gaming and recommendation system focuses on the five weakness areas that are vital aspects of lifelong educational skills, we coordinate with ScenicView with the desire to provide the support we can offer through our developed system.

During the first phase of the study, i.e., development of therapeutic games for adults with ASD, we coordinated with the teachers and therapists at ScenicView who offered their feedbacks on each of the therapeutic games designed for autistic adults. We presented multiple demo of the developed games to the administrators and program directors at ScenicView, incorporated the changes in each of the games as per their feedback, and completed the development of therapeutic games in all five weakness areas within an 18-month period which lasted from September 2017 to March 2019.

Upon developing the Minecraft therapeutic games that covered all the five areas, we began the testing phase, i.e., phase 2, by recommending therapeutic games and making the games accessible to the ScenicView students. The entire study was carried out from April 19, 2019 to July 19, 2019. We first installed the therapeutic games on the computer systems at ScenicView. With the help of teachers at ScenicView, we received a list of students who were willing to volunteer for the study. These students either played the games and filled up the questionnaire or simply filled up the questionnaire. A total of 21 students volunteered for the study. Prior to conducting the study, we scheduled a training session, where we invited all the 21 volunteer students to participate. We presented the questionnaire and the implemented therapeutic games. We also demonstrated and provided instructions on playing all of the therapeutic games. At the training session, we also asked the students to fill up the questionnaire for the very first time. The scores to the answers given by those students yielded their initial baseline scores which reflected the weakness areas prevalent in them.

After the training session, the volunteer students started playing the games recommended to them. Since a volunteer might not necessarily have weakness in all five areas, we determined the weakness areas prevalent in the student based on the student's initial baseline scores. Hereafter, according to the VideoGameGeek games (s)he has played in the past, we recommended to the student the games (s)he should play. A month after the student had played the recommended games the student was asked to fill up the questionnaire again. If the student had not crossed the threshold in a particular weakness area, i.e., the baseline score was less than 5.5, we recommended another game to the student in that area to be played according to our recommendation strategy. This process continued until either the student crossed the threshold of a designed weakness area or played the few games in each weakness area prevalent in the student during the 3-month period. The students were not clarified that their baseline score impacted the games recommended to them to ensure that they were not influenced while filling up the questionnaire.

7.2 The Data

The data for our performance evaluation includes the scores of the answers to the questionnaire provided by the volunteers at ScenicView Academy, in addition to the details of when and how long the volunteers played the games for. We collected our data for two different groups. The first group of students, which is called the *test group*, are volunteers who played the games recommended to them and filled up the questionnaire hereafter. The second group of students are categorized as the *control group*. Students in the control group did not play any of the therapeutic game but simply filled up the questionnaire at an interval of a 3- to 4-week period. The students were divided into the control and test group in order to study the effectiveness of our gaming and recommendation system by determining whether there were any differences in terms of enhancing their social-interactive skills by playing or

Table 7.1: Samples of baseline score pair for the Test and Control group

Test Group		**Control Group**	
Initial Baseline Score	Updated Baseline Score	Initial Baseline Score	Updated Baseline Score
4.6	5	5	5
3.8	3.9	3	3
4.5	4.5	3	2.6
3.4	4.2	4.2	5.5

not playing therapeutic games. Out of the 21 students from ScenicView who volunteered, 11 were interested in playing games and were assigned to the test group, whereas five were only interested in filling the questionnaire and were included in the control group, and the remaining five were not interested in either. For the volunteers in either group, we archived the baseline scores for each of the questionnaire filled up by them each month during the empirical study.

On a monthly basis, we observed the change in baseline scores achieved by the students in each weakness area with respect to the previous baseline scores and obtained pairs of baseline scores. We collected a total of 163 pairs of baseline scores in all weakness areas for the test group. We followed the same steps for the control group and obtained a total of 25 pairs of baseline scores. These pairs yield the data to evaluate the performance of our gaming and recommendation system using the significant test and samples of these pairs are shown in Table 7.1.

7.3 Wilcoxon Signed-Rank Test

Wilcoxon signed-rank test is a non-parametric statistical test that is used to compare pairs of sampled data. The design goal of the test is to study the difference between data pairs, since it is assumed that the difference of signs and magnitude of those pairs holds useful information. We use Wilcoxon rank test over other significant tests because it is robust as it is non-parametric, and the data does not need to be normally distributed.

Statistical significance indicates that the end result generated for an experiment is not merely due to a chance but there is a significant reason involved behind the result. The Wilcoxon test examines the null hypothesis over an alternative hypothesis. In statistics, *null hypothesis* is a theory that assumes the non-existence of statistical relationship or significance between two sets of observed data and is signified as H_0. *Alternative hypothesis*, on the other hand, is a proposed theory that contradicts the null hypothesis. It states that there is a statistically significant difference between two sets of observed data and the significance is due to some underlying reason rather than just by chance. Alternative hypothesis is often also called *research hypothesis* as it is developed by some study or research. The alternative hypothesis is represented as H_a or H_1.

In Wilcoxon test, rejecting or accepting the null hypothesis depends on the p-value. For a data sample that is large (i.e., with at least 25 pairs of data items), the magnitude of p-value determines if the null hypothesis should be accepted instead of the alternative hypothesis. The p-value is the probability that indicates whether the observed values are purely due to chance. Lower value of p means that the observed values are less likely due to chance which weakens the null hypothesis and is hence rejected. In general, a p-value less than 0.05 indicates that the pairs of values are statistically significant.

7.4 Performance Evaluation of Our Gaming and Recommendation System

For our conducted empirical study, the null hypothesis H_0 indicates that no significant difference exists between the pairs of baseline scores compiled by the surveys conducted by the volunteering students at ScenicView. The acceptance of null hypothesis indicates that our recommended games have no roles to play in the change of the volunteer's social-interactive skills. The alternative hypothesis H_1 signifies that the difference in the pairs of baseline scores is due to the recommended therapeutic games played by the students. The acceptance of alternative hypothesis over the null hypothesis implies that our gaming and recommendation system was able to recommend appropriate therapeutic games to the user

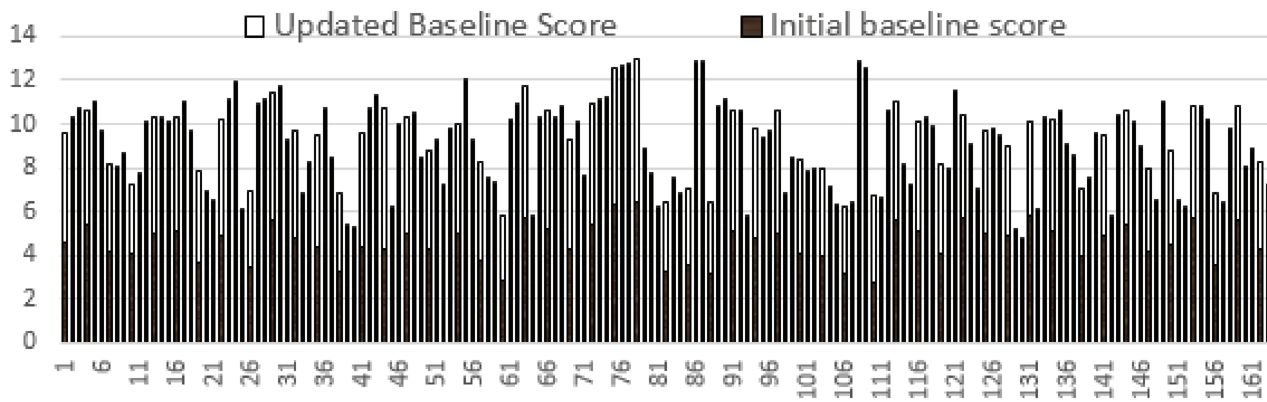


Figure 7.1: Distribution of baseline score pairs in the test group

Table 7.2: The Wilcoxon signed-rank test on the baseline scores of answers given by students in the test group

W-value	4239
Mean Difference	-0.87
Sum of Positive Ranks	4239
Sum of Negative Ranks	6639
p-value	**0.02034**
Z-value	-2.3205
Mean (W)	5439
Standard Deviation (W)	517.12

and the recommended games have an impact in enhancing the social-interactive skills of the students, which is the design goal of our gaming and recommendation system.

We ran the Wilcoxon test on the 163 pairs of baseline scores (and their distribution are shown in Figure 7.1) achieved by different volunteer students and the result of the test shows that the difference is *statistically significant*. The p-value is 0.02034, which is less than 0.05 and thus we draw the conclusion that H_1 is accepted. The acceptance of the alternative hypothesis proves that the changes in the scores for the students are due to the therapeutic games recommended to them and not just based on chance. The result of Wilcoxon test on the data is shown in Table 7.2.

In order to further strengthen the claim that the design goal of our gaming and recommendation system is achieved, we ran the Wilcoxon test again, but this time it was

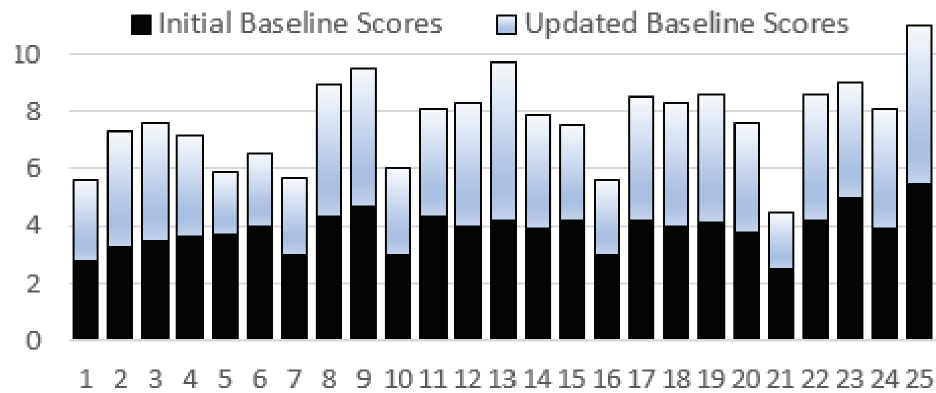


Figure 7.2: Distribution of baseline score pairs in the control group

Table 7.3: The Wilcoxon signed-rank test on the baseline scores of answers given by students in the control group

W-value	95.5
Mean Difference	-0.24
Sum of Positive Ranks	114.5
Sum of Negative Ranks	95.5
p-value	**0.72634**
Z-value	-0.3547
Mean (W)	105
Standard Deviation (W)	26.79

on the baseline score pairs achieved by the control group students (see the distribution of scores in Figure 7.2), and the outcome is shown in Table 7.3. The p-value for the test is 0.72634, which indicates that the difference between the data pairs is not *statistically significant*. Therefore, the null hypothesis for the control group is accepted which signifies that the change in baseline scores of the students in the control group is merely due to chance. This further strengthens our belief that the change of baseline scores for the test group is actually due to the recommended games played by the students and our gaming and recommendation system is effective.

```
=== Summary ===

Correlation coefficient                  0.9899
Mean absolute error                      0.1058
Root mean squared error                  0.1354
Relative absolute error                 14.2468 %
Root relative squared error             15.1158 %
```

Figure 7.3: Results of using KNN on the data of ScenicView students

7.5 Performance Evaluation of Our Automated Recommendation Strategy

During the 2^{nd} phase of our game recommendation methodology, we run KNN on the updated profile of a new user to obtain the affinity value for each therapeutic game. We employ KNN on the data obtained through the empirical study conducted on the volunteered students at ScenicView academy. We use the features mentioned in Table 6.1, set the number of neighbors, i.e., K, as five and run the KNN algorithm to obtain the affinity values. The results of the experiment based on 80-20 split of the training and test set are shown in Figure 7.3. (The training and test sets include the baseline scores of students at ScenicView academy.)

As seen in Figure 7.3, the accuracy of our automated game recommendation strategy is around 85%. The root mean squared error (RMSE) is 0.1354. The results indicate that KNN is effective in determining the *affinity values* for new users to automate game recommendation.

In order to experimentally verify that KNN is the appropriate choice among different existing regression and classification approaches, we run the same dataset through other approaches as well. We calculate the RMSE values for Multilayer Perceptron, Linear Regression, KNN, and SVM (Support Vector Machine). As shown in Figure 7.4, KNN has the least RMSE value, making it a suitable approach to be adopted.

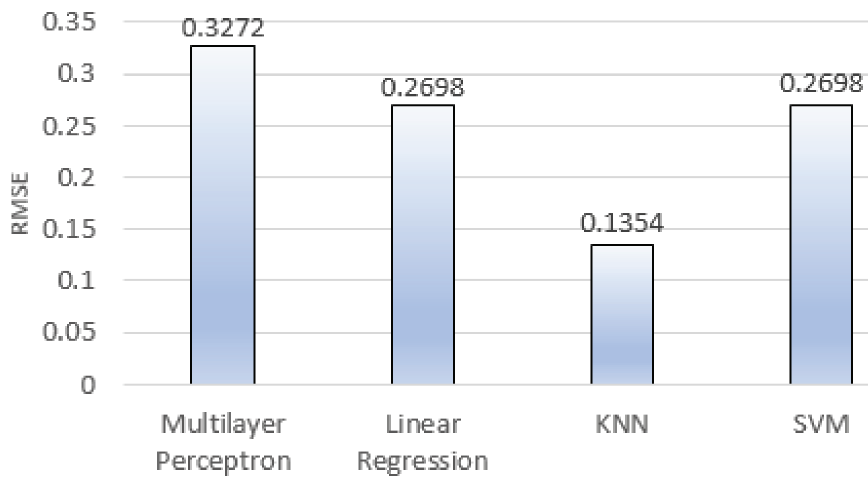


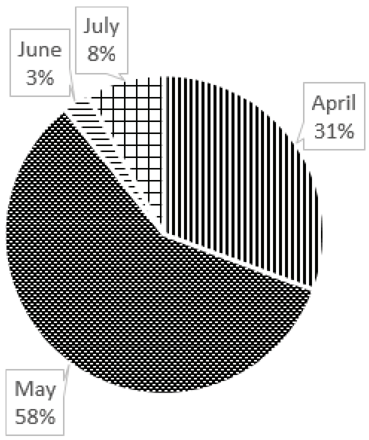
Figure 7.4: Performance evaluation of KNN over other approaches

7.6 Statistical Data of the Empirical Study

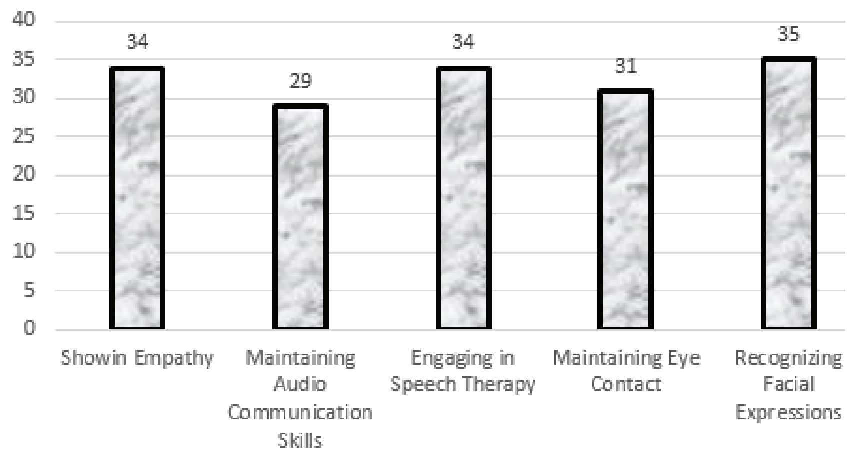
Apart from studying the effectiveness of our gaming and recommendation system, the data obtained from the test and control group also provide useful demographic information. Statistical data, such as the number of games that were played, number of players, types of game played, etc., convey important information. We studied on the distribution of games based on different metrices, such as the time when the games were played, the number of responses received in each weakness areas, and the gender of the students.

7.6.1 Distribution of Weakness Areas by Months

We studied the distribution of games played between April and July 2019 by the students. The distribution is represented as a pie chart in Figure 7.5(a). According to the data shown in the figure, the greatest number of games were played in the month of May as participants spent April in getting accustomed to playing the games and caught the pace in the following month.



(a) Distribution of games by months



(b) Distribution of games by weakness areas

Figure 7.5: Distribution of games played by participants by months and by weakness areas

7.6.2 Distribution of Weakness Areas by Volunteer's Responses

We also computed the distribution of responses of the students in each weakness area to determine the most prominent weakness area. Figure 7.5(b) shows that the greatest number of responses were given in 'Recognizing Facial Expressions', making it the most prominent weakness area among the participating students, whereas 'Developing audio communication skills' is the least prominent as there are least number of responses.

7.6.3 Distribution of Weakness Areas by Genders

Out of the total 16 participating students in the study, there were 9 males and 7 females. Figure 7.6 shows the distribution of weakness areas prevalent on the participants based on genders. The figure reflects that greater number of females struggle in 'Recognizing Facial Expressions', 'Developing Audio Communication Skills', and 'Engaging in Speech Therapy'. On the other hand, majority of males have deficiency in 'Maintaining Eye Contact', 'Recognizing Facial Expressions', and 'Developing Audio Communication Skills'. Based on these statistical data, we draw the conclusion that, 'Recognizing Facial Expressions', and 'Developing Audio Communication Skills' are the most common weakness areas irrespective of gender. Although Figure 7.6 shows that the number of responses in 'Developing Audio

4X Strategy, Action, Action Adventure, Action RPG, Adventure, Arcade, Augmented Reality, Beat 'em up, Classic Games, Clicker, Dating sim, Dungeon Crawler, Educational, Endless runner, Fighting, First person shooter, Fitness, Flight simulator, Hidden object, Interactive movie, Life simulation, Light gun shooter, Mana gement, Maze, MMO, MOBA, Other, Party, Pinball, Platform, Point-and-click, Puzzle, Racing, Real time strategy, Rhythm, Rouguelike, RPG, Run-n-gun, Sandbox, Scrolling, Shoot 'em up, Shooter, Simulation, Sports, Stealth, Strategy, Text adventure, Tower defense, Trivia, Visual novel, Walking simulator

Figure 7.6: Distribution of weakness areas based on genders

Communication Skills' is the least among all the weakness areas, the number of both male and female participants, whose baseline scores are lesser than 5.5 in this area, is greater than other areas, making it one of the most common weakness areas.

Chapter 8

Conclusions

Video games can be employed to improve the social-interactive skills of adults with ASD as the latter are fascinated with playing games. We have developed therapeutic games using Minecraft and recommended those games to adults with ASD based on their weakness areas. The effectiveness of our gaming and recommendation system has been verified through an empirical study which shows that the social-interactive skills of adults with ASD do improve by playing the therapeutic games recommended by our system.

Our research work contributes to the community of autistic adults. Typically, games in Minecraft are developed only for fun, but games that we have developed in Minecraft are both fun and therapeutic. Our recommendation system suggests suitable therapeutic games to adults with ASD. The proposed gaming and recommendation system is the first attempt to integrate therapeutic game development and recommendation system for skill enhancement of adults with ASD so that they can live a normal life and have a better future.

The suggestion of therapeutic games using our recommendation system is either semi-automated (i.e., the users fill up a questionnaire and based on that they are recommended therapeutic games to play) or is fully automated (i.e., a regression algorithm is used to determine suitable games for the users) to enhance the social-interactive skills of the users seamlessly. Without relying on user's medical profiles, which are confidential, and the required clinical studies, which are costly and labor-intensive, our gaming and recommendation system is self-sufficient and cost-effective. Also, our gaming and recommendation system personalizes the therapeutic game suggestion based on the weakness area of each individual, which is elegant and unique by itself.

The proposed gaming system attempts to close the gap between educational and entertainment games for (autistic) adults. Also, our therapeutic games are specifically focused towards adults with ASD, instead of children, which targets a more mature group of autistic patients. Moreover, the recommendation system is designed based on game attributes to determine suitable games for autistic adults who have deficiency in different weakness areas, which is significantly different from the traditional clinical therapy approach.

In the future, we would like to extend our gaming and recommendation system to other cognitive disabilities as well. As the gaming and recommendation system has achieved promising results for adults with ASD, the system could play a significant role across multiple disability areas as well. Furthermore, we plan to work with treatment centers for autistic adults, such as ScenicView Academy, to establish a regular course to enhance the social-interactive skills of adults with autism based on our gaming and recommendation approach that targets participant's individual needs, leading to mastery of essential skills efficiently.

References

[1] Autism and Developmental Disabilities Monitoring (ADDM) Network. Community Report on Autism, Centers for Disease Control and Prevention, 2018. https://www.cdc.gov/ncbddd/autism/addm.html.

[2] K. Bennett, R. Ramasamy, and T. Honsberger. The Effects of Covert Audio Coaching on Teaching Clerical Skills to Adolescents with Autism Spectrum Disorder. *Austim and Developmental Disorders*, 43(3):585–593, March 2013.

[3] G. Comrack, C. Clarke, and S. Buttcher. Reciprocal Rank Fusion Outperforms Condorcet and Individual Rank Learning Methods. In *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 758–759, 2009.

[4] Eye contact: Dont make these mistakes. https://www.canr.msu.edu/news/eye_contact_dont_make_these_mistakes.

[5] R. Dunifon and L. Gill. Games and Children's Brains: What is the Latest Research? Department of Policy Analysis and Management, College of Human Ecology, 2013. https://www.human.cornell.edu/sites/default/files/PAM/ Parenting/FINAL-Video-Game-Research-Brief-5.pdf.

[6] Edutainment Games. https://tvtropes.org/pmwiki/pmwiki.php/Main/EdutainmentGame.

[7] M. Garbade. Regression Versus Classification Machine Learning: What's the Difference?, 2018. https://medium.com/quick-code/regression-versus-classification-machine-learning-whats-the-difference-345c56dd15f7.

[8] C. Grossard, O. Grynszpan, S. Serret, A. Jouen, K. Bailly, and D. Cohen. Serious games to teach social interactions and emotions to individuals with autism spectrum disorders (asd). *Computers Education*, 113:195–211, 2017.

[9] A. Hiniker, D. Wong, and H. Williamson. Go Go Games: Therapeutic Video Games for Children with Autism Spectrum Disorders. In *Proceedings of ACM International Conference Proceeding Series*, pages 463–466, 2013.

[10] E. Johnson. Where Are All the Educational Video Games for Adults?, April 2015. https://www.recode.net/2015/4/24/11561860/where-are-all-the-educational-video-games-for-adults.

[11] J. Kirriemuir and A. Mcfarlane. Literature Review in Games and Learning. Technical report, A NESTA Futurelab, 2004. Report 8.

[12] Antti Knutas, Jouni Ikonen, Uolevi Nikula, and Jari Porras. Increasing collaborative communications in a programming course with gamification: a case study. In *Proceedings of the 15th International Conference on Computer Systems and Technologies*, pages 370–377. ACM, 2014.

[13] C. Manning, P. Raghavan, and H. Schutze. *Introduction to Information Retrieval*. Cambridge, 2008.

[14] M. Mazurek, P. Shattuck, M. Wagner, and B. Cooper. Prevalence and Correlates of Screen-based Media Use Among Youths with Autism Spectrum Disorders. *Autism and Developmental Disorders*, 42(8):1757–1767, August 2012.

[15] M. Mazurek, C. Engelhardt, and K. Clark. Video Games from the Perspective of Adults with Autism Spectrum Disorder. *Computers in Human Behavior*, 51(Part A):122–130, October 2015.

[16] A. Memari, N. Panahi, E. Ranjbar, P. Moshayedi, M. Shafei, R. Kordi, and V. Ziaee. Children with Autism Spectrum Disorder and Patterns of Participation in Daily Physical and Play Activities. *Neurology Research International*, pages 1–7, 2015.

[17] M. Namratha and T. Prajwala. A Comprehensive Overview of Clustering Algorithms in Pattern Recognition. *IOSR Journal of Computer Engineering*, 4(6):23–30, Sep-Oct 2012.

[18] How Video Games Benefit Students With Special Needs. https://www.aane.org/video-games-benefit-students-special-needs/.

[19] Y.-K. Ng and M. Pera. Recommending Social-Interactive Games for Adults with Autism Spectrum Disorders (ASD). In *Proceedings of ACM Conference Series on Recommender Systems (RecSys)*, pages 209–213, 2018.

[20] H.A.M Noor, F. Shahbodin, and N.C. Pee. Serious game for autism children: Review of literature. 2012.

[21] Parents of Children With ASD Generally Support Video-Game Play. https://leader.pubs.asha.org/doi/10.1044/leader.RIB1.20052015.12.

[22] P. Ohring. Web-Based Multi-Player Games to Encourage Flexibility and Social Interaction in High-Functioning Children with Autism Spectrum Disorder. In *Proceedings of International Conference on Interaction Design and Children, Vol. 7*, pages 171–172, 2008.

[23] L. Pashapoor, L. Kashani-Vahid, and E. Hakimirad. Effectiveness of cognitive computer games on attention span of students with intellectual disability. In *2018 2nd National and 1st International Digital Games Research Conference: Trends, Technologies, and Applications (DGRC)*, pages 82–87, 2018.

[24] N. Peñ-Miguel and S. Mx́imo. Educational Games for Learning. *Universal Journal of Educational Research*, 2(3):230–238, 2014.

[25] Research Autism. Introduction to the Issues Facing People on the Autism Spectrum, June 2018. http://www.researchautism.net/autism-issues/autism-issues-introduction.

[26] K. Ringland, C. Wolf, L. Boyd, M. Baldwin, and G. Hayes. Would You Be Mine: Appropriating Minecraft as an Assistive Technology for Youth with Autism. In *Proceedings of International ACM SIGACCESS Conference*, pages 33–41, 2016.

[27] G. Silva, A. Raposo, and M. Suplino. PAR: A Collaborative Game for Multitouch Tabletop to Support Social Interaction of Users with Autism. *Procedia Computer Science*, 27:84–93, 2014.

[28] J. Stewart. *Bridges We Build: The Art of Making Friends, A Guide for Connecting, Relating, and Socializing with Others.* ScenicView Academy, 2016.

[29] J. Torrente, Á. del Blanco, P. Moreno-Ger, and B. Fernández-Manjón. Designing Serious Games for Adult Students with Cognitive Disabilities. In *Proceedings of International Conference on Neural Information Processing*, pages 603–610, 2012.

[30] E. Vallefuoco, C. Bravaccio, and A. Pepino. Serious Games in Autism Spectrum Disorder An Example of Personalised Design. In *Proceedings of International Conference on Computer Supported Education*, pages 567–572, 2017.