Data Streaming and Real-time Analytics

Case study: Game Analytics
paper: A game analytics model to identify player profiles in single-player games

Outline

- Game Industry
- Game analytics
- Bartle’s Taxonomy

Game Industry

- A highly competitive industry
- Game designers need better strategies to create an attractive gameplay, high replay value (keep players coming back multiple times)
- Well-known strategy is to use game data to analyse player behavior, focusing on improving specific gameplay characteristics
    - Game Analytics

Game Analytics

- The science of online analysis and metrics of games
- Focusing on the use of player behavior data to increase revenue and avoid users leaving the game too early.
- Behavior analysis

Game Analytics

Metric	Definition & Use
ARPU (Average Revenue Per User)	Total number of unique players divided by total revenue in a given period. Can gauge the general business health of your game.
ARPPU (Average Revenue per Paying User)	Like ARPU but limited to purchasing players. Useful for gauging monetization strategies, such as in-game store design or store price points.
Unique Logins	Fairly self-explanatory, used to monitor active player population.
Conversion Rate	Total number of players who made purchases divided by unique logins. Gauges how well you convert free users into paid users.
Retention	Total number of players that logged into your game during two specific date ranges. Common gaps between reporting periods are 1-day, 7-day, and 30-day measures.
Avg. Session Length	Another self-explanatory metric, used to gauge how long your gameplay loop is fun or keeps players' attention.
Session Frequency	Number of players logged in, divided by the total number of login events. Gauges how often players engage with the game.
LTV (Lifetime Value)	Total number of unique players divided by total revenue generated. Gauges how high your ROI (return on investment) is per player.
Errors	Counts errors generated by code. Gauges how stable your game is.
Content Logs	Counts occurrences of specific logs unique to your game's content being triggered. Gauges popularity, stability, and engagement of specific content.

Behavior analysis

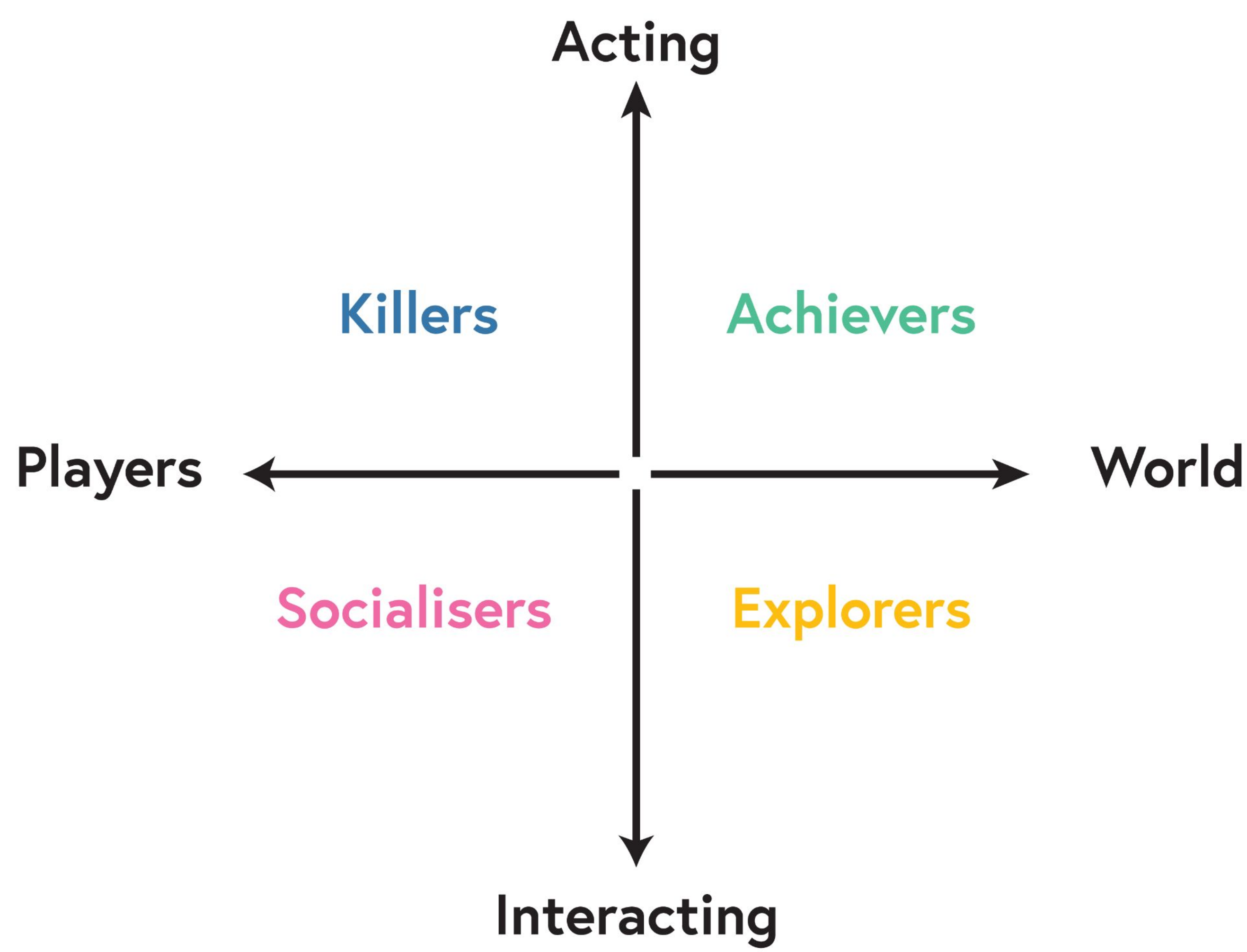
- Most of the works on player behavior analysis use **Robert Bartle's Taxonomy**
-
- The Bartle taxonomy of player types is a **classification of video game players** (gamers) based on a 1996 paper by Richard Bartle according to their preferred actions within the game.

1) Achievers

2) Killers

3) Socializers

4) Explorers



Bartle's Taxonomy

1. Achievers

- focused on mastering the game, on the rewards it has to offer. They share the world with other players, or non-playable characters (NPCs), and add a competitive element to the environment. Therefore, they are proud of their status in the game hierarchy, and how fast they reached their current level.

2. Killers

- focused on acting on other players, or NPCs, most of the time showing their superiority over them. They seek more power and abilities, that can help them affect others. Therefore, they are proud of their level of authority and their fighting skills.

Bartle’s Taxonomy

3. Socializers

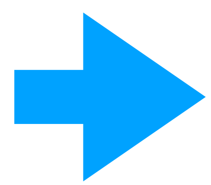
- focused on interacting and talking with other players, or NPCs. Also, finding more about other people is more interesting for socializers than competing, or bossing them. Therefore, they are proud of the relationships and of their influence towards other players.

4. Explorers

- focused on interacting with the world, the game environment. The sense of discovery or finding new areas and game elements fulfills them more than just achieving a great status in the game. Therefore, they are proud of their knowledge and of searching for new places and possibilities.

Goal


CURRENT
CLASSIFICATION:
CASUAL KILLER
COINS
5
730
SCORE
LIFE
HIGHSCORE
4620




CURRENT
CLASSIFICATION:
CASUAL ACHIEVER
COINS
18
1540
SCORE
LIFE
HIGHSCORE
4620
GAME OVER
YOU'RE A
CASUAL ACHIEVER
SCORE: 1540
COINS: 18
HIGHSCORE: 4620

Issue

- Most of the works use **unsupervised techniques, K-means**
  - to determine which player type fits better for a specific user, based on his/her mapped characteristics
  - However, these approaches usually get data from an already finished game, not having an algorithm ready for improvements in mechanics or difficulty.

Solution

- A combination of **K-means** + **Decision Tree algorithms** + **Bartle Taxonomy**

Solution

- A combination of **K-means** + **Decision Tree algorithms** + **Bartle Taxonomy**
  - Want to label each centroid, applying known achetypes to them.
  - to make this classification update regularly during the gameplay
  - This approach allows us to change the game mechanics and difficulty, having new centroids and decision trees for each newly added session
  - Allow programmers and designers to understand the classification steps quickly, e.g., which attribute was more decisive for each group.

Bartle’s Taxonomy

- In this work (a single player shoot’em up games)
  - Achievers, who focus on collecting items and coins
  - Killers, who focus on killing enemies
- Combined with
  - Casual
  - Hardcore

Dataset

A0) Number of direction changes (Mean);

A1) Position in X axis (Mean);

A2) Position in Y axis (Mean);

A3) Total time in movement (Mean);

A4) Number of items collected (Total);

A5) Number of coins collected (Total);

A6) Number of destroyed enemies (Total);

A7) Percentage of game completed (Total);

A8) Number of shots (Mean);

A9) Number of shots on target/enemies (Mean);

A10) Number of shots without enemies (Mean);

A11) Number of shots taken or

Number of lives lost (Total);

We define twelve attributes that are updated every half a second (0.5 second is the time interval).

Algorithm

1. Applied K-means clustering with K=4

2. Assigned C0, C1, C2, C3 to each instance.

3. Applied Decision Tree

Decision Tree

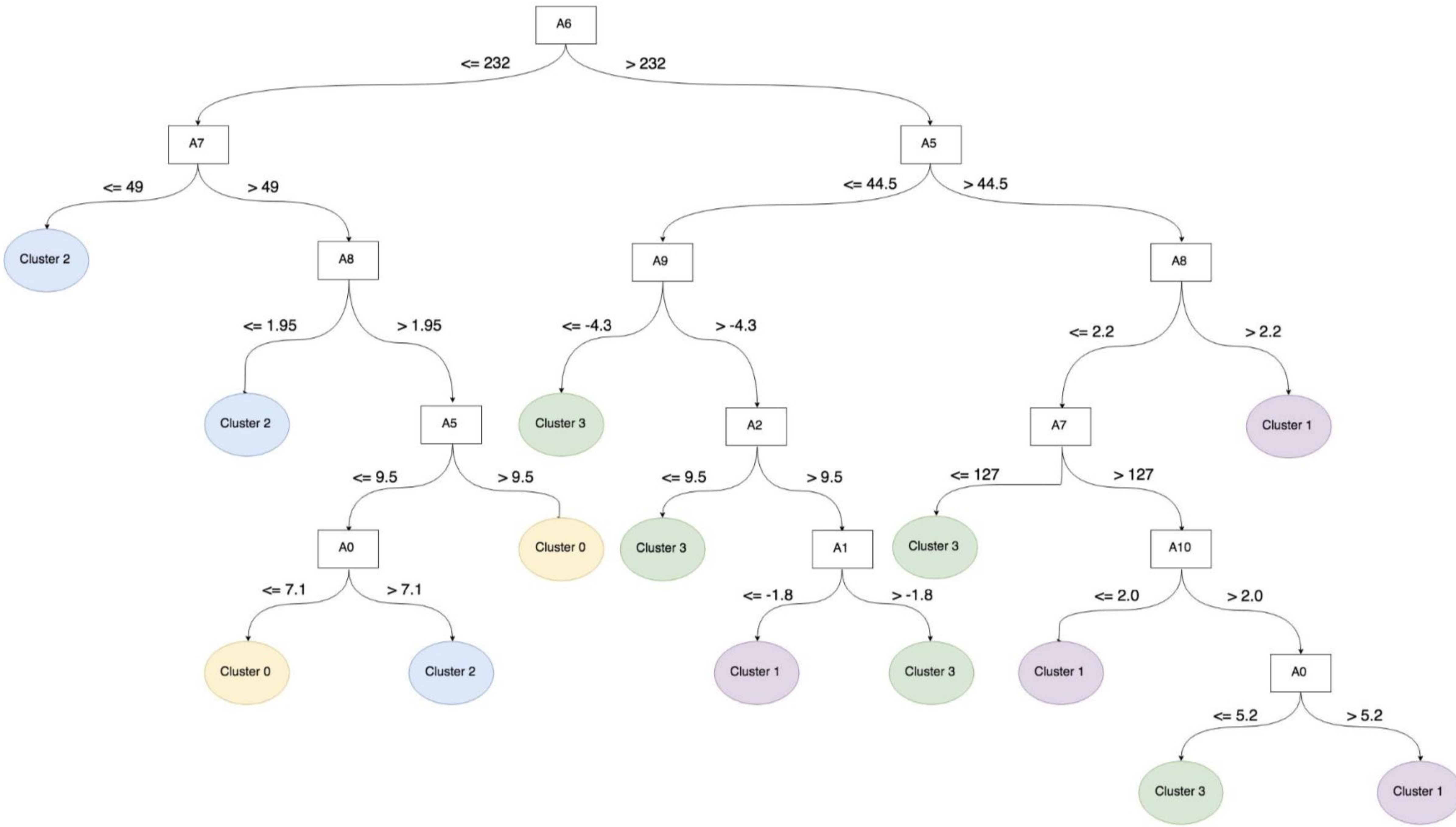


Figure 6. Decision Tree generated by *Accord.NET* algorithm

Algorithm

1. Applied K-means clustering with K=4

2. Assigned C0, C1, C2, C3 to each instance.

3. Applied Decision Tree

4. This paper select A5, and A6

5. Now we could iterate in the centroids array [C0,C1,C2,C3], searching which one of them had the bigger A5 value. This first one would be the "**Hardcore Achiever**" centroid. Then we would search, among those remained, which one had the bigger A6 value. This second one would be the "**Hardcore Killer**" centroid. This same process was repeated for the rest of the array, choosing, respectively, the "**Casual Achiever**" and the "**Casual Killer**" centroids.

To cross check with real player's profile

Questionnaire Design

- We designed two questionnaires.
- The first verifies if the player is classified as an **Achiever** or a **Killer**. It was based on the work by Schneider et al. (2016), which presents a questionnaire containing twenty questions, resulting in a percentage for each player type.
- Not use binary questions forcing the player to fit in a profile. They use, instead, the same five answers for every question:
    - "I do not understand/I do not identify myself" (0 points);
    - "I identify myself a little" (1 point);
    - "I identify myself partially" (2 points);
    - "I identify myself" (3 points);
    - "I identify myself totally" (4 points).

Questionnaire Design

Achiever

- “I like to conquer new badges in games”;
- “I get impressed with players that conquered high rewards”;
- “I play electronic games until the end with 100% of achievements”;
- “I love new items and medals”;
- “I like exposing my achievements (for example, on Facebook)”.

Questionnaire Design

Killer

- “I am very competitive in games”;
- “I like exploding things in games”;
- “My favorite games are first person shooters”;
- “I am known for my aggressiveness in games”;
- “I do not like talking in games, what I really like is shooting”.

Achiever or Killer

To decide whether the player is an Achiever or a Killer, we decided to sum the points related to the questions of each archetype, and get the maximum value from their result, as shown in equation 3. If the sum result is equal for both types, the player is classified as both, lowing the chances of the game classification being wrong. This also happens, for instance, if the player is defined as 55% Killer and 45% Achiever, i.e. he/she is classified as both if the distance between both Killer and Achiever percentage is below or equal to 10 percentage points.

$$PT = \max\left(\sum_{i=1}^{5} A_i, \sum_{j=1}^{5} A_j\right), (A_i, A_j) \in [0, 4] \qquad (3)$$

Casual or Hardcore

- "I always deal with technology and seek for new releases and trends" (7 points);
- "I like to have the latest high-end computers/consoles" (7 points);
- "I'm willing to pay anything for a game" (5 points);
- "I prefer violent/action games" (1 points);
- "I prefer games that have depth and complexity" (3 points);
- "I play games over many long sessions" (10 points);
- "I always search for the game industry latest information" (6 points);
- "I frequently talk about games, both via social media and with people" (10 points);
- "I always feel happy when completing (or defeating) a game" (7 points);
- "I don't get easily frustrated while playing a game" (9 points);
- "I am usually engaged in competition with myself, the game, and other players" (6 points);
- "I started playing games when I was little" (2 points);
- "I have played all the types of game genres, and I constantly compare one game to another" (10 points);
- "I buy games and consoles on their pre-release, or import them from other countries to be one of the first to play" (9 points);
- "I think of modifying and extending some of the games I play" (8 points);

Casual or Hardcore

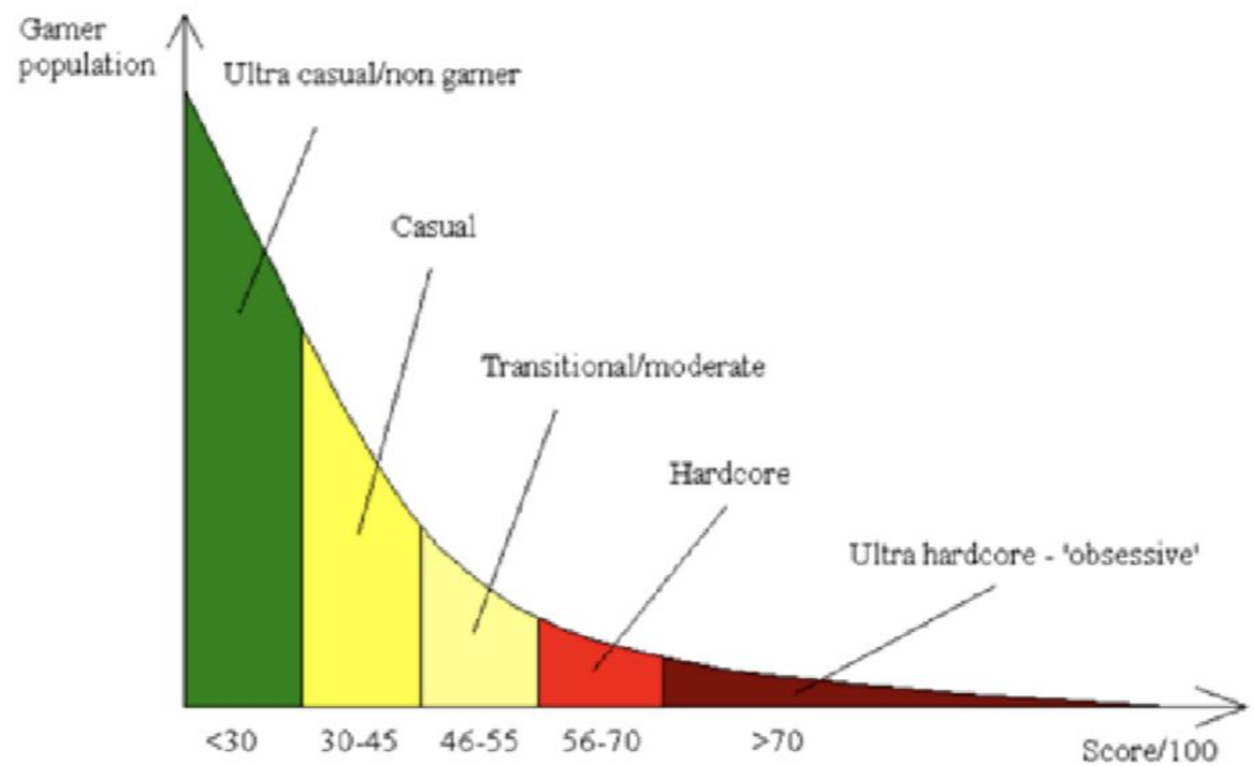


Figure 7. Casual and Core by gamer dedication

1) Casual gamer - Has GD factor below or equal to 45%;
2) Moderate gamer - Has GD factor between 45% and 55%, with these limits included;
3) Hardcore gamer - Has GD factor above 55%.

$$GD = \frac{\sum_{i=1}^{15} A_i \times Q^i}{\sum_{i=1}^{15} 5 \times Q^i} \tag{4}$$

Results

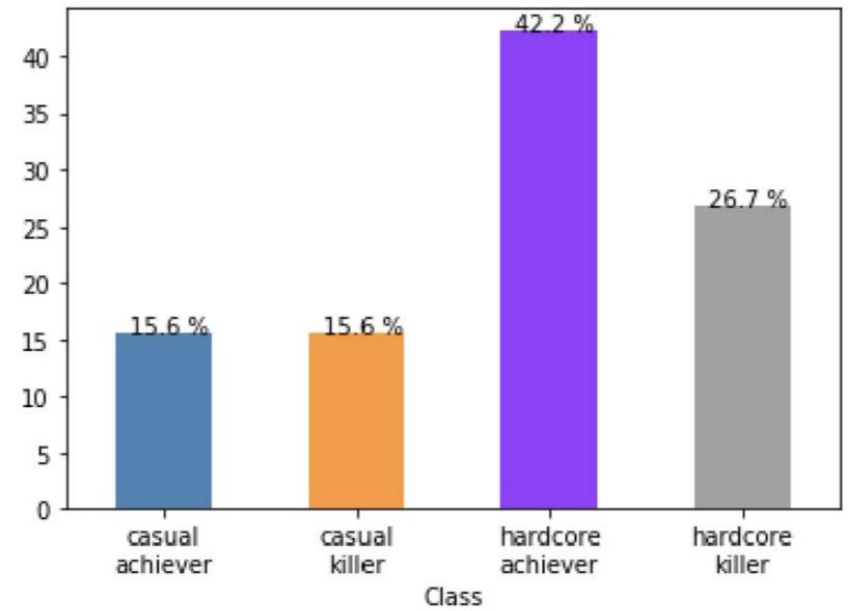


Figure 8. Total of each archetype found on game sessions

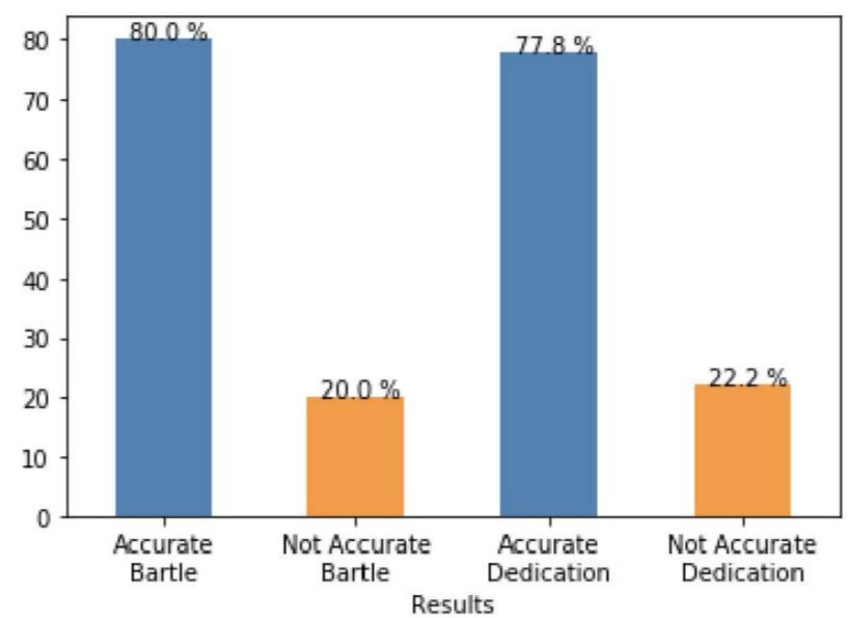


Figure 9. Accuracy results for both dedication and Bartle archetypes

Future works

Our approach allows the creation of an adaptive model, which would base itself on the player classification to vary the difficulty of the game, changing parameters like

- enemy speed,
- the number of enemies on screen,
- enemy fire rate,
- player fire rate, and
- other relevant gameplay characteristics.

There are many future works to do. We should investigate if player type scores do significantly predict player experience.