Alex Fletcher | Nick Ballou | Sebastian Berns
Feb 2020
IGGI Methods and Data 2

Responsible Innovation Framework for Win-rate Prediction in Clash Royale

1. Brief Summary of Current Research

The topic of the present research study was the ability to predict the winner of a given match in Clash Royale, given each players' ranks and decks. Clash Royale is a free-to-play (freemium) mobile game for iOS and Android. In this real-time strategy game, players build a deck from any 8 of 70 total units (cards), and then place these cards at various locations into the arena. The units will march toward the opposing player's three towers, fighting each other along the way; whichever player destroys all three of their opponent's towers (or has destroyed the most towers when time ends) wins the game.

To train the prediction models, we used a database of 723,742 1v1 matches that took place between July 12, 2017 and July 31, 2017 and were publicly posted to Kaggle[1] after being scraped from the Clash Royale API. After processing and removing matches from alternative modes or special events and draws, we were left with a final sample of 387,997 matches (775,994 decks) from 81,864 unique players. The data for each match includes each player's deck, the card levels from 1-13 associated with each of the player's 8 units, the result of the match including towers destroyed, and the players' in-game usernames and clans.

This data was then used to provide illuminating descriptive statistics about the win rate (i.e., when this card appears in a deck, what percentage of the time does that deck win) and use rate (i.e., in what percentage of decks does this card appear) of particular cards, and to train machine learning models to predict the outcome of the match before it happens. The primary goal of the present research was to provide players with new insights into their performance, including how to improve deck selection and understand favorable vs. unfavorable matchups. The work may also have positive benefits for esports casting, balancing strategies for developers, and player experience; these anticipated outcomes, and others, are further detailed in Section 2.1 below.

Below we report our responsible innovation plan. We begin in section 2 with aspects of responsible innovation that we successfully integrated during this two-week project. In section 3, we describe deviations from our ideal responsible innovation plan and shortcomings of the work. Finally, in section 4, we propose rough guidelines and a timeline to ensure responsible innovation in future studies in this area.

2. Responsible Innovation in Current Research

In this section we describe to what extent and how our research fits the EPSRC's formal responsible innovation framework. We discuss the AREA guidelines'[2] four-fold approach to anticipate, reflect, engage and act. The identified stakeholders include the game's developers, its competitive and casual players, as well as anyone with a general or a specialized interest in the game, such as game researchers, commentators, and general esports audience.

2.1 Anticipate

Anticipation in the AREA framework refers to "describing and analysing the impacts, intended or otherwise [...] that might arise. This does not seek to predict but rather to support an exploration of possible impacts and implications that may otherwise remain uncovered and little discussed." In preparing and motivating the current study, the research team discussed various possible applications, implications,

[1] https://www.kaggle.com/s1m0n38/clash-royale-matches-dataset
[2] https://epsrc.ukri.org/research/framework/area/

and consequences of the work. We divide these here into positive (i.e., beneficial to society in some way) and negative (i.e., harmful) outcomes, noting of course that these are not mutually exclusive.

2.1.1 Positive outcomes

The primary positive outcome we envisioned was the ability to offer **new and actionable insights** for players looking to improve their gameplay. In Clash Royale, strategically selecting the eight cards from your collection so as to maximize synergy, effectively counter common opponent strategies, and align with the player's strengths is a crucial part of success. Having access to additional information about individual cards' strengths and weaknesses, synergistic combinations, and distributions of card usage and effectiveness. These types of data analysis could also assist developers in **balancing** cards in the meta-game, as well as offering a degree of quantification regarding the balance of luck and skill in the game. At a macro-level, both the insights available to assist players in improving their performance and the fine-tuning of the game's balance could lead to greater engagement, enjoyment and motivation. Additionally, as Clash Royale has a rich esports community, this research could have an impact to improve the quality of casting (commentary), whereby the prediction system could be used as leverage for detailed discussion on strategy of the players.

2.1.2 Negative outcomes

In addition to the positive outcomes identified above, the team also discussed possible negative consequences. The first of these was identified in relation to data privacy. The data used for this study is pre-existing and publicly available; in this data, players' in-game names and clans are shown. While these are typically not directly identifiable (unless a player enters FirstnameLastname as their tag, for example), some in-game names and clans can give indications about a player's name, country of origin, or age (e.g., an "AmericanLegends" clan, or a player named Jimmy2004). These player names can also be used to search public databases like StatsRoyale.com for a player's detailed match history and card levels. There is further potential for these statistics to be linked with some indirect proxies for spending behavior. Because Clash Royale is a freemium game, leveling up one's cards without paying requires consistent play over a long period of time. By looking at a player's match history in combination with the level of their cards, one could make inferences about a player's spending; if a player has high card levels but has played relatively few matches, they are likely to have purchased premium gems in order to speed up the leveling process.

Other possible negative outcomes were identified with regard to facilitation of gambling both inside and outside of the game. Outside of the game, a highly accurate prediction model could make betting on esports tournaments more appealing or profitable for bookmakers. Inside of the game, increased engagement could lead indirectly to increased in-game spending on the game's loot boxes, which have been linked in previous research to problematic gambling in both adolescents and adults. We noted that Clash Royale is rated acceptable for ages 9+ on the Apple App Store, which means that these gambling-like features are available to children.

2.2 Reflect

Reflection refers to the process of thinking critically about the purposes of and motivations for the research, alongside its potentially unexamined assumptions or framings. We noted early on that this research relies to a certain extent on a particular motivational model of video game use and of players; by focusing on winning and strategic factors that lead to in-game success, a variety of other motives like social connection, escapism, or audiovisual stimulation may be ignored. This does not mean that the research goals are not useful; it does, however, mean that the positive outcomes will only be relevant to a particular subset of more competitive players. This was something we carefully considered.

Our assumptions about the usefulness of our proposed research does also highlight an area of ignorance. As we discuss in section 3.3 below, we have not been in direct contact with professional or casual players and thus operate purely based on our personal understanding of what might be interesting and helpful to our target beneficiaries of Clash Royale players.

One further area in which we noticed a number of implicit assumptions was in the interpretation of a model fit to our data. An accurate prediction model suggests impartial objectivity, certainty and confidence in its prediction. But a high test score on a data set of a toy problem that is several years old does not necessarily imply a model is robust enough to be employed in a real-world setting. Its limitations (in the generalisability and actual confidence of prediction) are not well understood by the general public. This might lead someone to believe such a model lowers the inherent risk in gambling on a match winner, for example. Having reflected on this particular issue, we would like to minimize these negative possible implications. How this could be achieved is outlined in section 4 below.

2.3 Engage

With regard to engagement (i.e., "opening up [the above] visions, impacts and questioning to broader deliberation, dialogue, engagement and debate in an inclusive way"), the current study did so only through speculation. We spent time to consider who the stakeholders were and what feedback, criticisms, interests, or worries they might have, and attempted to integrate those perspectives in our work. We were not, however, able to communicate with these other parties. This shortcoming is discussed below in section 3.3.

One area of successful engagement worth noting was found in our consideration of data visualization. Figures can be superficial or misleading (Weissgerber et al. 2015), and this is compounded by the fact that our target audience consists of people who are largely not trained statisticians or academics. This was an important consideration when preparing slides and figures for dissemination.

2.4 Act

The process of thinking about and discussing the points in the sections above shaped the method we used in the study, and also influenced the particular research outputs. One important methodological decision we made in consideration of responsible innovation was model choice; given the explicit goal of providing players with actionable insights into their gameplay, we decided to use as many descriptive statistics as possible alongside maximally interpretable machine learning models. In this way, we do not simply predict outcomes with a black box algorithm, but rather are able to share meaningful information with players, e.g. the likely effect of adding a Fireball to one's deck when using Golem as well.

An illustrative example of how we acted in response to our responsibility for engagement concerns the data visualization challenges mentioned above. One important results figure from our work compares win rates for each individual card; in theory, these win rates may span from 0 to 1, but in practice, developers typically strive to keep cards within approximately .45 and .55. We debated how to most honestly reflect the importance of these differences (i.e., whether to crop the y-axis to a certain range which may give exaggerated impressions of size differences, or span the entire range from 0 to 1, which may give the impression of small/meaningless differences between cards). In the end, we decided to partially crop the y-axis, but to highlight that range verbally during our presentation—we would make sure to emphasize the scale, or include both a cropped and uncropped version, in future public dissemination of this information.

3. Deviations from and Shortcomings of Responsible Innovation Plan

3.1 Anticipate

With regard to anticipation, there were no significant deviations from our responsible innovation plan. The research team was able to ask the necessary 'what if' questions, consider contingencies, and consider what is known, likely, plausible, and possible. This process was well-timed so that it was early enough to be constructive but late enough to be meaningful.

3.2 Reflect

We identified no major shortcomings regarding the reflect aspect of our innovation plan. Throughout the study, the team continued to look critically at the work we were doing and any implicit assumptions. This is a primary effect of preparing this responsible innovation report in parallel to the research project.

3.3 Engage

Engagement was the framework aspect in which we felt we deviated most significantly from responsible innovation guidelines. Responsible innovation in the context of this work requires an open dialogue with the relevant stakeholders, which in this case consist primarily of the players and developers. Given the time and resource constraints of the project, it was not possible to discuss with either of these parties, and thus we were limited to 'imagining' our outcomes. While our research team includes players of Clash Royale, it is not sufficiently representative for the different stakeholder groups. By talking to players, we would have had the opportunity to learn about particular aspects of the game in which players feel they could improve or in which data may answer their existing questions. Crucial during this process would be to ensure that a representative sample of different gamers' opinions are taken into account, accounting for differences in (among other things) players' degree of engagement with the game, gender, race, national origin, and age. Talking to the developers, Supercell, would have improved our understanding of what game analytics professionals in their company have already done, what data they might be willing to share, their openness to possible collaboration, and avenues they may have identified as interesting areas for future work.

3.4 Act

While we were able to act in some ways based on our initial responsible innovation plan, some other ways were beyond the scope of the current project. As mentioned above, while we identified the need to talk to stakeholders, we were not able to do so. Neither players, developers, nor researchers working in related areas were consulted.

Data privacy serves as another example; while we were able to foresee possible problems regarding the anonymity of this data, due to the lack of time, resources, and expertise in data encryption, we were not able to take any additional steps to anonymize the data on our end. When considering best practices, we agreed that the most responsible method would be to collect data from the API oneself and anonymize it upon collection. Player names and clans can be replaced with a random identifier, and thus retain all the information in the data without risking the security of the player's data. More sophisticated encryption or blinding techniques like fully homomorphic encryption could also be used to ensure that data is protected in case of a breach, or that even the researchers never see the non-anonymized data.

4. Improvements for Future Work

To conclude, we propose a tentative timeline for similar research conducted either by ourselves or by others (see Figure 1 for a gantt chart representing the research timeline). The work should begin with a careful consideration of the study aims and its potential outcomes. These outcomes should then be discussed with players and developers in order to assess its feasibility, usefulness, and the possibility of other consequences not identified by the research team. Other researchers are also stakeholders in this; a review of the relevant literature will illuminate where the study fits within the field and the importance of its intended contributions. In response to any new findings, concerns or suggestions during this process, modifications to the study should be made. In this sense, the research design process constitutes a full cycle of the Anticipate, Reflect, Engage, and Act cycle—this process is ongoing, however, and should continue throughout the remainder of the work.

At this point, prior to data collection, a clear strategy for data collection, privacy protection, and data management should be agreed upon in written form. For example, where and when will the data be posted, if at all? Will the data be shared with other researchers, and if so, how will it be sent and will that require any additional anonymization? These questions should be answered before data collection begins. The data collection itself should then make sure to collect only the relevant data, adhering to the guidelines specified by the developer and/or third parties, and to anonymize that data as soon as it is collected.

The main quantitative work can be conducted in a similar way as the present research; however, with a longer time period, the analyses could grow in complexity and adapt to the idiosyncrasies of the data, leading to potentially more accurate models, a greater depth of insight, and ultimately greater benefit to players and developers. Once the work has concluded, special attention should be paid to accurate data visualization and reporting, and this reporting should be done transparently such that the work is maximally reproducible (Munafò et al. 2017); to the greatest extent possible, the researchers should endeavor for their published work to be open access, for code and materials to be openly shared, and for data to be shared for others to use as long as privacy concerns can be met.

However, acting upon the concerns discussed in section 2.2 about the general public's limited knowledge of the limitations of a machine learning model, publication of the pre-trained models should be done cautiously; potentially limiting public access and in any case including a detailed explanation of its basic assumptions and limitations.

The final part of the study involves a return to the Engage aspect of the AREA framework. Given that the purpose of this work was, among other things, to provide insights for players, future work should endeavor to make the analyses available in as convenient a way as possible. We have identified one possible way of doing so through a simple web app where players can enter their current cards and receive suggestions for complementary cards that yield the highest win rates, or enter in specific opponents' cards that the player would like to counter most effectively; the app will then calculate the card that will have the greatest positive change in win rate against that opponent deck. This app should again be created in consultation with users and designed in accordance with their needs.

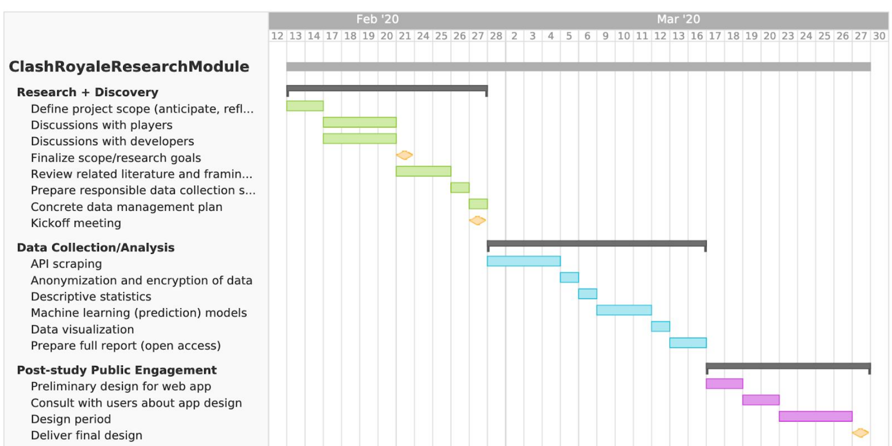

Fig. 1: Timeline of future responsible research

5. References

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. Nature Human Behaviour, 1(1), 0021. https://doi.org/10.1038/s41562-016-0021

Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm. PLOS Biology, 13(4), e1002128. https://doi.org/10.1371/journal.pbio.1002128