A study on Channel Popularity in Twitch

Ha Le[1], Junming Wu[2], Louis Yu[3], and Melissa Lynn[4]

[1] Gustavus Adolphus College, 56082 MN USA
hle5@gustavus.edu
[2] Gustavus Adolphus College, 56082 MN USA
jwu@gustavus.edu
[3] Gustavus Adolphus College, 56082 MN USA
lyu@gustavus.edu
[4] Gustavus Adolphus College, 56082 MN USA
mklynn@gustavus.edu

Abstract. In the past few decades, there has been an increasing need for Internet users to host real time events online and to share their experiences with live, interactive audiences. Online streaming services like Twitch have attracted millions of users to stream and to spectate. There have been few studies about the prediction of streamers' popularity on Twitch. In this paper, we look at potential factors that can contribute to the popularity of streamers. Streamer data was collected through consistent tracking using Twitch's API during a 4 weeks period. Each user's streaming information such as the number of current viewers and followers, the genre of the stream etc., were collected. From the results, we found that the frequency of streaming sessions, the types of content and the length of the streams are major factors in determining how much viewers and subscribers streamers can gain during sessions.

Keywords: Social Network Analysis · Twitch · Machine Learning.

1   Introduction

Social Media has made enormous impacts on how we share information and experiences with each other. There exists many social networking platforms that allow us to post images, videos, and other content. In the past few years, as the video games and e-sports industry grows rapidly, generating billions dollar of revenue from viewership [1], it has become clear that there is a need for players to share gaming experiences with one another. Since then, many gaming platforms have gained momentum in streamers and viewership. Among them Twitch has stood out as the best-known for streaming videos games and broadcasting major e-sports events.

Twitch provides users with a professional and unique streaming experience by giving streamers the opportunity to build strong, connected communities through customized dashboards. And, it gives viewers novel ways to support their favorite streamers through monthly subscriptions and other donations. There have been research on the popularity of users in social networking platforms

[18, 15]. Even though some research has attempted to analyze Twitch communities[12, 4], few has examined the factors that contribute to the popularity of Twitch channels and their corresponding streamers.

In this paper, we collect data by crawling to explore the distribution of popularity among streams. We categorize them and conduct analysis in order to identify characteristics of successful streamers. We have concluded that a successful streaming strategy includes: (1) streams do not exceed 5 hours, (2) at least 5 streams per week and (3) combine non-gaming and gaming content. Applying those characteristics, we also predict the growth rates of channels through machine learning classification model.

The rest of this paper is organized as follows. In Section 2, we provide some background information. In Section 3, we survey related work. Section 4 describes the setup of our experiments, followed by results in analysis in Section 5. Finally, we give our conclusion in Section 6.

2   Background

2.1   Social Media

Online social media has grown tremendously over the past few years. According to a study in 2016 [11], 68% of the adults in the United States use the Internet and 79% of them engage in at least one form of social networking platforms. Facebook, a major online social networking platform, announced that they have more than one billion users around the world in 2012, while LinedkIn reported to have half a billion users in 2021 [5]. Social media platforms have created online communities that allow people to make and share content, to make connections and to communicate with each other. There has been research on online communities's effects on our political [24], social [20, 27] and economical [6] believes.

2.2   Streaming Services

Streaming media is a booming industry. The main goal of these services is to deliver continuous and uninterrupted content to their users. Streaming platforms usually have two components: the component the broadcasts content and s comments/audience interaction Section. The streaming content on these platforms is diverse. Today major social media platforms such as Facebook, Instagram, YouTube, etc. have adopted streaming components. The most popular streaming services are for broadcasting movies and TV shows. This is a very competitive market with new platforms developed every year. For example, Netflix, HBO Max, Disney+, Amazon Prime Video, YouTube etc. Music streaming is also a large component of the industry, in which Spotify is the leading platform. Many people uses the streaming components in social networking platforms such as Facebook and Instagram for sharing experiences, interacting with followers, and sometime selling products and advertising.

2.3   Twitch

Twitch is a streaming service focusing primarily on video games and e-sports. In 2020, Twitch, as one of the largest streaming platforms, has attracted over 30 million average viewers daily.

One of the most recognizable features in Twitch is its ability for streamers to grow their communities and attract viewers based on their content. Streamers can also customize dashboards, enforcing regulations and administering bots. Twitch allows viewers to directly support their favorite streamers by buying and giving tokens. For avid viewers, they can choose to subscribe to their favorite streamers and receive corresponding perks provided by the channels. This allows streamers, from build their own communities, to receive financial support, which makes it possible for some to develop careers as professional streamers. Apart from direct support from the fans, Twitch also grant partnerships and affiliations which help streamers to generate incoming. Other than the streamers, the other important part of the Twitch community are the viewers. Viewers watch the streams and create dynamic, interactive environments by conversing with the streamers and with each other in the chatroom. Research has shown that people are able to develop meaningful, maintainable relationships through the Twitch platform [26].

3   Related Works

Studies on Streaming Services: Pires and Simon [22] presented a 3-months dataset on Youtube and Twitch. Their preliminary studies show that both platforms, especially Twitch, are able to generate rich services at any time. The popularity of the channel is more heterogenous than what have been observed in other social media platforms. Zhang et al. [32] conducted a survey on consumer behaviors on Spotify, a popular music streaming platform. They identified some key characteristics on user behaviors; this includes the users' favorite time of the day, the length of a successive user session and the tendencies to switch devices. Lee et al. [16] examined the impact of music streaming services on record sales in South Korea and showed that music streaming has a positive impact on the record sales. Surprisingly, the number of songs in an album and the number of singers in a group also have an impact on record sales as well.

Studies on Streaming Communities: Gandhi et al. [8] discussed how streamers make connections with their audiences. Todd and Melacon [29] explored the impact of streamers' genders on the perception of viewers in live streaming sessions and found that the likelihood of a viewer interacting with streamers depend heavily on the sex of the streamers. Cai et al. [2] looked at micro-communities in streaming platforms and showed that there are correlations between the imposed regulations, the vibe of the channels, and harassment tendencies within the community. Cai and Wohn [3] examined the role of volunteer moderators in streaming communities.

Studies on Twitch: Although we cannot find research similar to ours, there have been other studies on Twitch. Seering et al. [25] looked at chatbots, programs designed to perform automated tasks during streams, and showed how they can help communities grow and evolve. Poyane [23] looked at chat logs from Twitch channels to create a topic model of viewers discussions.

Twitch Users' Profiles: Many researchers have analyzed the profile of Twitch users. Kim et al. [14] looked at users' differences in sentiments through messages in chat; from that, they attempted to identify potential subscribers' characteristics. Lee [17] approached the same topic of whether a viewer will subscribe to a channel through machine learning, building models such as logistic regression, SVM, decision tree and random forest. Han and Sukhee [13] collected data to examine the distinctions between channels in South Korea. Zhao et al. [33] attempted to identify human factors such as openness, personal affordance and gaming levels that can contribute to the popularities of channels.

Predicting the Popularity of Users on YouTube, Twitter, and Spotify:
 While there are few research on the popularity of users on Twitch, there are studies on other streaming platforms such as YouTube, Twitter, and Spotify. Figueiredo, Benevenuto and Almeida [7] studied the growth patterns of video popularity on YouTube, Based on the data collected, they categorized the videos into three categories and observed the characteristics of videos that tend to motivate viewers into watching, thus contribute to channel growths. Pinto, Almeida, and Goncalves [21] present two simple models for predicting the popularity of Web content on YouTube based on historical information given by popularity measures. Ma, Yan and Chen [19] achieved fast prediction of long-term video popularity on YouTube. To understand the propagation of news Tweets, Wu and Shen [30] built a model that can predict the final number of retweets for a news tweet based on information collected from super nodes on Twitter. Gao et al. [9] studied popularity prediction of posts in microblogs. Hu and Sci [31] looked at predicting peak time popularity of tweets based on Twitter hashtags.

 Studies have been done on the prediction of user popularities on Spotify, one of the most popular music streaming platforms. Araujo, Cristo and Giusti [28] approached this problem as a classification task and employed classifiers built on past information from the platform's Top 50 Global ranking as well as acoustic features from the songs. They found that RBF kernel gave the best results with an AUC higher than 80% when predicting the popularity of songs two months in advance. Wu and Sun [10] adopted Principle Components Analysis (PCA) and model blend-in, avoiding significant error by high dimension, and gave a relatively precise prediction on the popularity of Spotify's top music.

4   Experiment Setup

Our data set consists of streaming information from 20,000 random active Twitch users, gathered using Twitch's API. Additional information on the games them-

selves was crawled from the web using a web crawler. The Twitch API does not provide achieved streaming information so all streams were real-time data obtained by consistently tracking activities from users over the course of two month. We start by generating a dataset of 10,000 active Twitch users by requesting 20,000 random streams currently on the platform and extracting the ids from the streamers.

Gathering Streaming Information: Over the course of 4 weeks, we closely monitor the activities of the selected accounts. Every week we choose 4,000 users and monitor their activities; every 30 minutes we record their current status and their number of followers. If the account is streaming, we record the id of the stream as well as the current number of viewers, the game tag and the streaming language. After one week, we switch to another set of 4,000 users and repeat the process.

Gathering Additional Information: Apart from the real-time information gathered from live streams on Twitch, we also looked for additional information related to the channel. Specifically, we gathered information on any additional videos and clips uploaded to the channels after the streams have finished. That is, after streams have ended, the users can upload videos cut from parts of the streams to their channel. Using the streams' ids, we were able to track them. If there were special moments during the stream that have the potential to go viral after, streamers can choose to cut a short clip from the stream and upload it as highlight. Using the streams' ids and the videos' ids, we were able to track the highlights.

We think that the popularity of games played during streams is also an important factor on the views. Using a web crawler, we obtained daily popularity of each game title using Twitch's self-reported statistics.

Over the course of one month, we were able to get the complete data of 10,058 users, with more than 55000 streams and 50000 video clips. Our dataset consists of streams in 20 different languages, with both gaming and non-gaming contents.

5   Analysis

5.1   The Power Law Distributions

First, we look at the popularity of streamers in the dataset by examining the distributions of their corresponding number of followers and viewers. Figure 1 shows the distribution of followers (a), the distribution of average viewers during streams (b), and the distribution of viewers tuning in at the start of the stream (c). All three follow the power law with exponents 1.25, 1.11 and 2. This indicates an underlying relationship between the popularity of streamers and their ability to attract followers/viewers. That is, streamers' popularity will grow at a rate proportional to his/her current followers/viewers count.
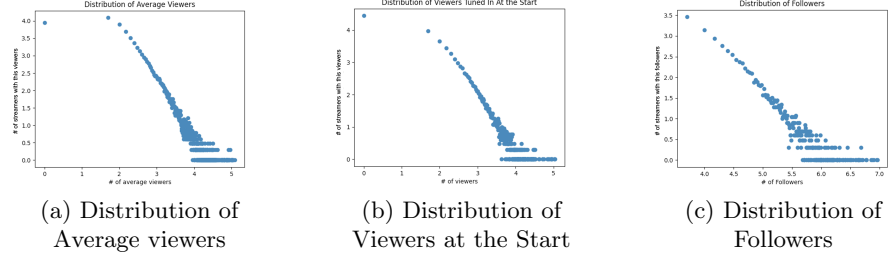
(a) Distribution of
Average viewers

(b) Distribution of
Viewers at the Start

(c) Distribution of
Followers

Fig. 1: Power Law Distributions

5.2   Distribution of Followers Gained and Bot Detection

Next, we looked at the average distribution of followers gained during the week. Users' activities during the week are categorized as active and inactive. Figure 2 shows the distribution of followers gained during streaming (active) period, and the distribution of followers gained during non-streaming (inactive) period. We observed that the graphs are spiky. They also suggest that at times streamers were getting more followers during inactive periods than during active periods. These results are curious to us. A closer look at the data reveals several suspicious surges and declines in the number of followers for many streamers. We suspect bot interference from these abnormal behaviors. To remedy the above problem,
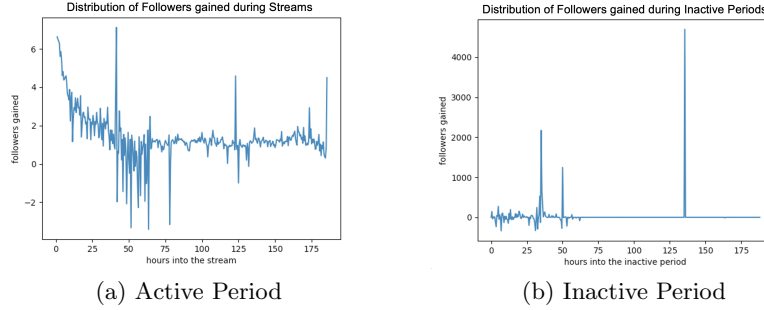

(a) Active Period

(b) Inactive Period

Fig. 2: Distribution of Followers Gained (Before)

we employed a simple, naive algorithm to detect bot behaviors and delete users with such from our dataset. Our algorithm works by looking for sudden rise and drop in the number of followers between each 30 minutes timeframe. We calculated the absolute differences in followers between one time frame and the next. If the difference between timeframe i and $i+1$ is at least 100 times larger than the difference between timeframe i and $i-1$ and timeframe $i+1$ and $i+2$ for user A, we mark such user as abnormal and delete user A from the dataset.

After running the algorithm on our set, we detected 226 users who have shown abnormal growth or decline during the inspected week, among which 110 streamers have more than 10,000 followers, and 67 of them have more than

100,000 followers. We provide a pseudocode of the algorithm below in Algorithm 1.

Algorithm 1: Bot Bahavior Detection

Input: streamers dataset
Output: streamers dataset w/o bot behavior
for *userA in streamers dataset* **do**
    for *timeframe[i] in userA* **do**
        if $|timeframe[i+1] - timeframe[i]| >$
        $100 \cdot |timeframe[i] - timeframe[i-1]|$ **and**
        $|timeframe[i+1] - timeframe[i]| >$
        $100 \cdot |timeframe[i+2] - timeframe[i+2]|$ **then**
            mark userA as suspicious
    if *userA is suspicious* **then**
        delete userA from the dataset
return new dataset

Even though heuristic algorithm is incomplete and inaccurate, it significantly helps us eschewing outliers and suspicious behaviors that tainted the results of the experiment. The result dataset after the application of bot detection shows major improvement on the distribution of followers growth during active and inactive period, which is shown in Figure 2.

The distribution of followers growth during active and inactive periods after applying our bot detection process is shown in Figure 3. We observe that despite the occasional spikes in graph (b), there are not many fluctuations in the number of followers during inactive periods, which implies stability in the number of followers for most streamers when not streaming. A closer look at graph 3(a) reveals that most streamers experience the majority of increase in popularity during the first couple of hours of the streams; such increase starts to taper off when the streams go longer than 20 hours.
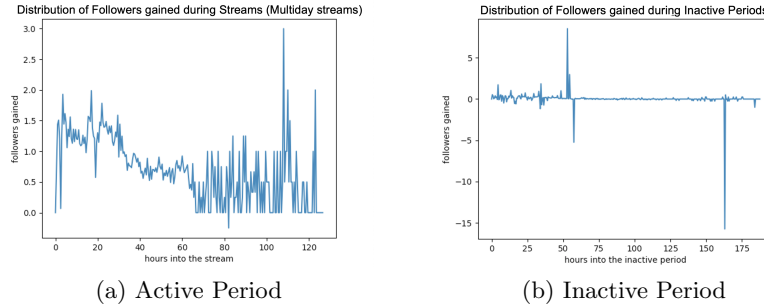

(a) Active Period   (b) Inactive Period

Fig. 3: Distribution of Followers Gained (After)

5.3   Categorization of Streamers

Section 5.1 has shown that streamers' popularity can affect their ability to gain more followers/viewers. To identify other factors that can contribute to the success of a channel, we categorize the streamers in our dataset by their initial number of followers (the number of followers at the start of our observation period) into 4 categories (see Table 1).

Table 1: Categories of Users

Categories	Initial number of followers	Number of users
small	0-5000	1857
medium	5000-10000	1177
big	10000-100000	3383
mega	> 100000	1062

We mostly focus on small and medium streamers in our dataset since they are less susceptible to the rich gets richer effect discussed in Section 5.1. In the following Sections, we look at how streaming hours and the duration of streams can affect streamers' ability to gain more followers during active periods.

Duration of the Streams Figure 4 shows that the duration of a stream can affect the number of followers gained per streaming hour. For most users streaming less than 6 hours often brings the most optimal gain in followers per hour, while streams that last more than a day usually has less followers growth per hour.
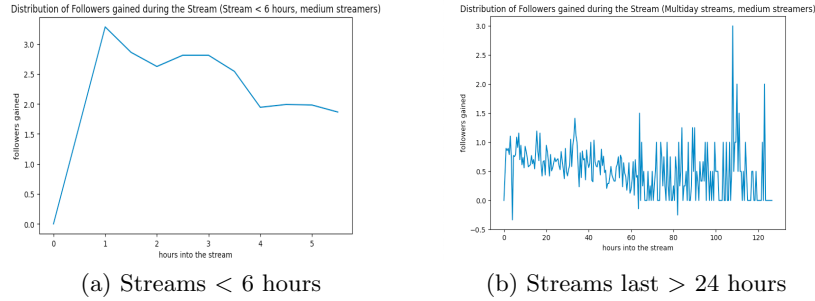

(a) Streams < 6 hours   (b) Streams last > 24 hours

Fig. 4: Distribution of Followers Gained

Streaming hours Since streamers live and stream in different time zones, it is difficult to identify a specific optimal streaming hours for all users. Instead, we will identify users by their streaming languages and look for the best streaming time for each set. After analyzing the average followers gained for each 30 minutes timeframe, we noticed some trends. In particular, streams in some languages have several peak hours where streamers on average gain simnifically more followers,

while streams in other languages only ave one peak hour. Figure 5 shows an example for each (English and Spanish). We will use UTC timezone for both graphs in Figure 5.

From Figure 5 we observe that streamers broadcasting in Spanish gain the most followers on average between 13:00 - 14:00 UTC time, while streamers broadcasting in English has several peak hours: 5:00 - 7:00, 16:00 - 17:00, and 20:00 - 21:00 UTC time.
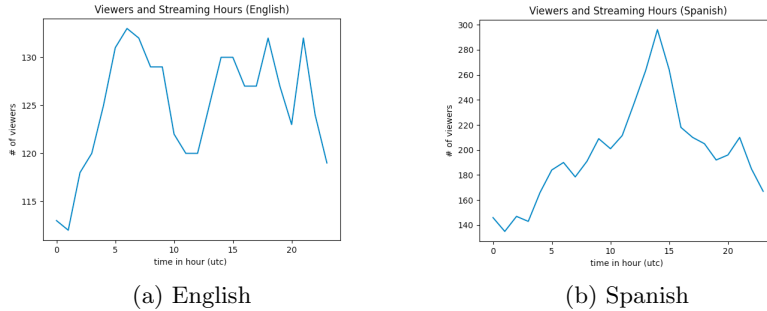

(a) English   (b) Spanish

Fig. 5: Optimal Streaming Hours for Different Language

5.4 Isolating Outliers

In this Section, we employ a method to identify the best strategy for streamers to attract followers, outliers isolation. By identifying streamers who perform better than their peers, we hope to understand what they have in common in streaming characteristics.

First, we only consider streamers with initial number of followers between 0 and 20000. We then separate them into 20 "buckets" based on their initial popularity (the number of followers recorded at the start of our observation period): streamers with initially 0 - 1,000 followers, streamers with initially 1,000 - 2,000 followers, streamers with initially 2,000 - 3,000 followers, and so on, for up to streamers with initially 19,000 - 20,000 followers. For each bucket, we identify streamers who performed better than their peers by isolating the top 5% who gained the most subscribers. From the 5738 streamers being considered, we were able to identify 198 outliers. Next, we investigate what these outliers have in common and describe the result in the next few Sections.

Language Distribution Figure 6(a) shows the pie chart of the distribution of streaming languages among the outliers. The most popular language used is English, followed by Spanish. This is to be expected since the most used languages among Twitch streamers are English and Spanish. We observe from the chart that all languages have at least one representative in the outliers. Thus, we conclude that languages used by streamers have little or no affect on channels' growth rate.
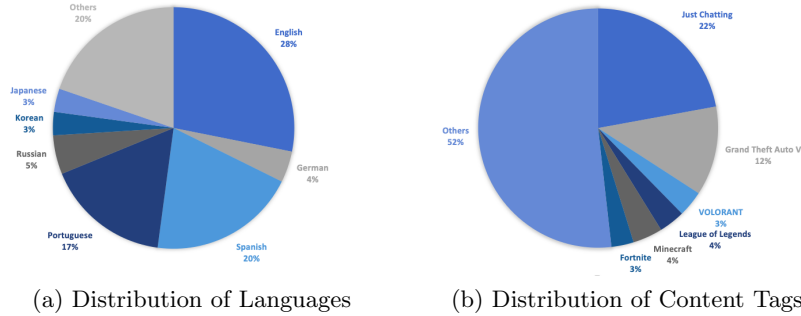
(a) Distribution of Languages   (b) Distribution of Content Tags

Fig. 6: Distribution of Used Languages and Content Tags for outliers

Content Distribution Figure 6(b) shows the pie chart for the types of gaming content streamed by outliers. Surprisingly, the most prominent type of content is Just Chatting, a non-gaming related tag, followed by the more well known gaming tags such as League of Legends, Grand Theft Auto V, and Minecraft. For a further investigation, we divided tags into two categories: gaming and non-gaming. Table 2 shows the percentages of gaming and non-gaming tags used by the outliers. We observe that 50% of the streams by the outliers have non-gaming tags, while only 29.22% of the streams in our entire dataset have non-gaming tags. This shows that successful streamers are more likely to stream a mix of non-gaming and gaming content.

Table 2: Content Distribution

	entire dataset(%)	outliers(%)
gaming	70.78	50.00
non-gaming	29.22	50.00

Length of the Stream Figure 7(a) and 7(b) show the distribution of the length of the streams for outliers and the entire dataset. We observe that most streams from the outliers generally stream much shorter streams in comparison to other streamers, with majority of the streams lasting less than 2 hours. For the entire population in our dataset, the majority of the streams last up to 6 hours, with many streams lasting for days.

Streaming Frequency Figure 7(b) and 7(c) show the distribution of streaming frequency for the outliers and for the entire dataset. We observe that most outliers stream more than 5 times per week, with many streaming between 5 - 8 times per week.

Overall, we observe from our results that a streaming strategy with multiple short streams and mixed content is the most effective.

From the information above, we identify a common strategy of attracting more followers/viewers: 1) stream at least 5 times per week, 2) stream no more than 2 hours in each streaming, 3) gaming and non-gaming contents are mixed

(a) Distribution of the Length of
the Streams for the Entire Dataset

(b) Distribution of the Length of
the Streams for the outliers

(c) Distribution of the # of
Streams for the Entire Dataset

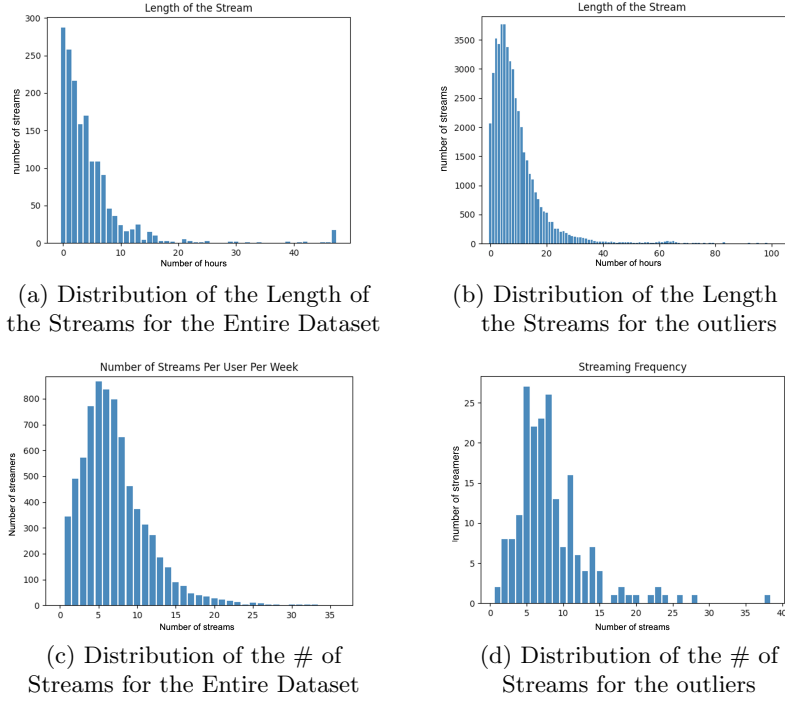(d) Distribution of the # of
Streams for the outliers

Fig. 7: Distribution of Length of the Streams and Streaming Frequency

and evenly distributed as streaming contents, 4) stream more than 40 hours per week.

Table 3 shows the statistics comparison between streamers apply that strategy and entire dataset. On average, streamers with that strategy gained triple as much followers as entire dataset.

Table 3: Comparison of followers gained

	Entire dataset	Streamers fit the strategy
Average Follower Gained	101	342
Median Follower Gained	29	67
Average Viewers per Hour	101	101

5.5   Predicting the Growth of a Channel

Our prediction task involves classifying users with initially 0 - 20k followers into classes based on their growth rate during the week. We aim to use users' quantitative streaming parameters to identify "good", "average", and "bad" channels (in reference to the channels' ability to attract followers).

In our case, we chose to use a random forest classification model. Random forest classifier is a commonly used learning algorithm. Its mechanism involves the creation of multiple decision trees for training and the use of majority voting

for the prediction process. For our study, we used the implementation provided in the Python Scikit Learn package.

First, we classify channels based on their growth rates. We reused the bucket methods described in Section 5.4 and divided the channels into 20 buckets based on their initial number of followers. Then, if a channel looses followers during the observation process, we consider it a *bad* channel. If a channel gained more followers than the average followers gained by channels in its bucket, we consider it a *good* channel. Otherwise, the channel is marked as *average.* Table 3 shows the average followers gained by channels in each bucket.

Table 4: Average followers gained for channels in each bucket

Initial followers range	Average followers gained	Initial followers range	Average followers gained
0-1000	33.39	10000-11000	164.40
1000-2000	61.74	11000-12000	205.43
2000-3000	77.60	12000-13000	207.54
3000-4000	76.16	13000-14000	133.64
4000-5000	92.0	14000-15000	214.60
5000-6000	117.29	15000-16000	309.73
6000-7000	143.46	16000-17000	188.78
7000-8000	139.39	17000-18000	127.06
8000-9000	159.08	18000-19000	359.18
9000-10000	155.12	19000-20000	197.55

Table 4 shows the size of each class. We can see that the classes are unbalanced. To improve on the accuracy of our model, we used oversampling to partially fix the imbalance of our dataset. We used Python Imblearn library's implementation of Synthetic Minority Oversampling Technique (SMOTE). We selected 80% of our data (4590 users) as the training set, and held 20% (1148 users) as the testing set. We apply oversampling to the training set to get a balanced dataset. Then, the oversampled training set is fed to our random forest model to learn. This model is then applied to the testing set to verify its accuracy. We give the parameter fed to the model as well as the response variable

Table 5: Classes and class sizes

Class	Number of users
good	1339
average	4213
bad	186
total	5738

the model returned. The learning parameters are: the initial number of followers and their bucket, the number of streams per week, the number of streams with less than 5 hours in length, the number of streams with 5 - 10 hours in length, the number of streams with more than 10 hours in length, gaming/non-gaming content (1 = yes, 0 = no), whether the streamers streamed popular games (1 = yes, 0 = no). They response variable is a prediction of whether the channel will perform good, bad or average (1 = good, 0 = average, -1 = bad).

Lastly, we report on the accuracy of our trained model. We show the average accuracy, precision, recall and F1 measures of the model for 8 different maximum depths of the trees below.

Table 6: Result of the model

Max depth	Accuracy(%)	Precision	Recall	f1
2	40.06	0.61	0.40	0.43
4	54.61	0.63	0.54	0.58
6	66.63	0.65	0.66	0.65
8	71.34	0.66	0.71	0.67
10	71.51	0.66	0.71	0.67
12	70.73	0.65	0.70	0.66
14	68.55	0.63	0.68	0.64
16	66.81	0.62	0.66	0.64

From Table 6, we can see that the model with maximum trees' depth $= 10$ performs best with accuracy upto 71.55%.

6   Conclusions

Over the past few years, streaming services have become an emerging source of media content. With the surge in popularity of e-sports and competitive gaming, Twitch has grappled the opportunity to become a major player in the gaming market. It hosts thousands of major gaming events and millions of independent streamers every year. With a customizable platform, it allows streamers and viewers to build their own communities and provide financial support for streamers.

In this paper, we evaluate the different factors that could potentially contribute to streamers' success on Twitch. Our hope is to shed some light on the complex, intricate ecosystem of Twitch communities.

Our data was collected using Twitch's API. We tracked the weekly activities of 20000 active Twitch users over the course of one month and gathered over 3M data points, with more than 55,00 streams and 50,000 videos, We inspected the features of streamers and their corresponding streams and identified which their importance. This includes the growth in followers, the number of views per video, the number of views per stream, and the content of the streams.

Our model provides a tool for both Twitch communities and streamers. For Twitch communities, our model provides a guide to evaluate streamers and help identify future streaming celebrities. For streamers, we provide a self-evaluation method to help them with improving their streaming characteristics.

References

1. Borowy, M., et al.: Pioneering esport: the experience economy and the marketing of early 1980s arcade gaming contests. International Journal of Communication **7**, 21 (2013)

2. Cai, J., Guanlao, C., Wohn, D.Y.: Understanding rules in live streaming micro communities on twitch. In: Conference on Interactive Media Experiences (IMX'21) (2021)
3. Cai, J., Wohn, D.Y., Almoqbel, M.: Moderation visibility: Mapping the strategies of volunteer moderators in live streaming micro communities (2021)
4. Churchill, B.C., Xu, W.: The modem nation: A first study on twitch.tv social structure and player/game relationships. In: 2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom). pp. 223–228 (2016). https://doi.org/10.1109/BDCloud-SocialCom-SustainCom.2016.43
5. Dewing, M.: Social media: An introduction, vol. 1. Library of Parliament Ottawa (2010)
6. Edosomwan, S., Prakasan, S.K., Kouame, D., Watson, J., Seymour, T.: The history of social media and its impact on business. Journal of Applied Management and entrepreneurship **16**(3), 79–91 (2011)
7. Figueiredo, F., Benevenuto, F., Almeida, J.M.: The tube over time: Characterizing popularity growth of youtube videos. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining. p. 745–754. WSDM '11, Association for Computing Machinery, New York, NY, USA (2011), https://doi.org/10.1145/1935826.1935925
8. Gandhi, R., Cook, C.L., LaMastra, N., Uttarapong, J., Wohn, D.Y.: An exploration of mental health discussions in live streaming gaming communities. Frontiers in psychology **12**, 751 (2021)
9. Gao, X., Cao, Z., Li, S., Yao, B., Chen, G., Tang, S.: Taxonomy and evaluation for microblog popularity prediction. ACM Transactions on Knowledge Discovery from Data (TKDD) **13**(2), 1–40 (2019)
10. Ge, Y., Wu, J., Sun, Y.: Popularity prediction of music based on factor extraction and model blending. In: 2020 2nd International Conference on Economic Management and Model Engineering (ICEMME). pp. 1062–1065 (2020)
11. Greenwood, S., Perrin, A., Duggan, M.: Social media update 2016. Pew Research Center **11**(2), 1–18 (2016)
12. Hamilton, W.A., Garretson, O., Kerne, A.: Streaming on twitch: Fostering participatory communities of play within live mixed media. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. p. 1315–1324. CHI '14, Association for Computing Machinery, New York, NY, USA (2014). https://doi.org/10.1145/2556288.2557048, https://doi.org/10.1145/2556288.2557048
13. Han, S.: Characteristics and comparison of popular channels on internet game broadcasting: Focus on twitch tv. The Journal of The Institute of Internet, Broadcasting and Communication **16**(4), 7–14 (2016)
14. Kim, J., Bae, K., Park, E., del Pobil, A.P.: Who will subscribe to my streaming channel? the case of twitch. p. 247–251. Association for Computing Machinery (2019), https://doi.org/10.1145/3311957.3359470
15. Kong, Q., Rizoiu, M.A., Wu, S., Xie, L.: Will this video go viral: Explaining and predicting the popularity of youtube videos. In: Companion Proceedings of the The Web Conference 2018. p. 175–178. WWW '18, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE (2018). https://doi.org/10.1145/3184558.3186972, https://doi.org/10.1145/3184558.3186972

16. Lee, M., Choi, H., Cho, D., Lee, H.: Cannibalizing or complementing?. the impact of online streaming services on music record sales. Procedia Computer Science **91**, 662–671 (2016)
17. Lee, S.H.: A study of prediction of live-stream subscriptions: the case of twitch (2019)
18. Ma, C., Yan, Z., Chen, C.W.: LARM: A Lifetime Aware Regression Model for Predicting YouTube Video Popularity, p. 467–476. Association for Computing Machinery, New York, NY, USA (2017), https://doi.org/10.1145/3132847.3132997
19. Ma, C., Yan, Z., Chen, C.W.: LARM: A Lifetime Aware Regression Model for Predicting YouTube Video Popularity, p. 467–476. Association for Computing Machinery, New York, NY, USA (2017), https://doi.org/10.1145/3132847.3132997
20. O'Keeffe, G.S., Clarke-Pearson, K., et al.: The impact of social media on children, adolescents, and families. Pediatrics **127**(4), 800–804 (2011)
21. Pinto, H., Almeida, J.M., Gonçalves, M.A.: Using early view patterns to predict the popularity of youtube videos. In: Proceedings of the Sixth ACM International Conference on Web Search and Data Mining. p. 365–374. WSDM '13, Association for Computing Machinery, New York, NY, USA (2013), https://doi.org/10.1145/2433396.2433443
22. Pires, K., Simon, G.: Youtube live and twitch: a tour of user-generated live streaming systems. In: Proceedings of the 6th ACM multimedia systems conference. pp. 225–230 (2015)
23. Poyane, R.: Toxic communication on twitch. tv. effect of a streamer. In: International Conference on Digital Transformation and Global Society. pp. 414–421. Springer (2019)
24. Qin, B., Strömberg, D., Wu, Y.: The political economy of social media in china. Tech. rep., Working paper (2016)
25. Seering, J., Flores, J.P., Savage, S., Hammer, J.: The social roles of bots: evaluating impact of bots on discussions in online communities. Proceedings of the ACM on Human-Computer Interaction **2**(CSCW), 1–29 (2018)
26. Sheng, J.T., Kairam, S.R.: From virtual strangers to irl friends: Relationship development in livestreaming communities on twitch. Proc. ACM Hum.-Comput. Interact. **4**(CSCW2) (Oct 2020). https://doi.org/10.1145/3415165, https://doi.org/10.1145/3415165
27. Siddiqui, S., Singh, T., et al.: Social media its impact with positive and negative aspects. International Journal of Computer Applications Technology and Research **5**(2), 71–75 (2016)
28. Soares Araujo, C.V., Pinheiro de Cristo, M.A., Giusti, R.: Predicting music popularity using music charts. In: 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA). pp. 859–864 (2019)
29. Todd, P.R., Melancon, J.: Gender and live-streaming: source credibility and motivation. Journal of Research in Interactive Marketing (2018)
30. Wu, B., Shen, H.: Analyzing and predicting news popularity on twitter. International Journal of Information Management **35**(6), 702–711 (2015)
31. Yu, H., Hu, Y., Shi, P.: A prediction method of peak time popularity based on twitter hashtags. IEEE Access **8**, 61453–61461 (2020)
32. Zhang, B., Kreitz, G., Isaksson, M., Ubillos, J., Urdaneta, G., Pouwelse, J.A., Epema, D.: Understanding user behavior in spotify. In: 2013 Proceedings IEEE INFOCOM. pp. 220–224. IEEE (2013)
33. Zhao, K., Hu, Y., Hong, K., Westland, J.: Understanding characteristics of popular streamers on live streaming platforms: Evidence from twitch.tv. Journal of the Association for Information Systems (11 2020)