FC BARCELONA

A Simplified Approach To Building a Video Recommendation System on FC Barcelona's Official App

Tomás Ravalli; Hector Aurélio Assumpçao Palácios; Víctor González Faría

Abstract

The level of user interaction on digital products can be influenced by the relevance of the content that are presented to the user. In sports entertainment businesses such as FC Barcelona, video is the main form of content consumption, such as game analysis and highlights, preview of upcoming games, press conferences, and original content in documentary / docuseries format, to name a few. The selection of the videos presented to the user in each interaction then plays an important role in the level of interaction and retention of the users as well as in the conversion of occasional users into registered users and finally into paying users. The club has implemented an out-of-the-box video recommendation system with predetermined parameters that has reached an effectiveness of 30% of successful views in February 2021. With this baseline in mind, this article describes a way of classifying users and videos, analyzes the results observed by user and video cluster, and the percentage of successes of those interactions, to finally provide a recommendations matrix with the objective of increasing the total interaction of users in the app and ultimately converting logged users into paying users.

Keywords: digital business; mobile app; video content; recommender system

Introduction

FC Barcelona has designed its main app as a means of attracting and retaining users globally, eventually increasing its fan base and its income through different revenue streams. A combination of static content (related to the structure and history of the Club) and dynamic content (related to the multiple sporting events in which the Club participates) are presented to the user in each interaction, and in addition to photos, tables and texts, the videos play an essential role in the time and level of user interaction, presenting opportunities to extend the interaction time through the consecutive viewing of multiple videos, if they are of interest to the user.

The club has implemented a basic out-of-the-box recommendation system offered by one of its technology providers, and after having collected enough information on its performance, it hopes to be able to contribute to the recommendation algorithm in a way that helps it achieve maximization objectives. the level and time of interactions, with a view to converting users to the paid level.

Following industry standards [1], in this work a unique interaction metric will be used, based on the percentage of the length of the video that is viewed by the user. If the displayed length is more than 30% the user is considered to have had a successful interaction and it is assumed that presenting the user with this type of material will lead to more successful views.

This work proposes a matrix recommendation methodology, based on the classification of videos and users, and the analysis of the interactions observed in a given period. Through this, a knowledge matrix is built that indicates what kind of video is more attractive for each kind of user, making this information available to the algorithm to enhance the level and time of future interactions.

This methodology is based on the concepts presented in works such as [2] and allows establishing a starting point for a mechanism of recommendations that can also be improved in the future, as more information is collected that feeds it.

After presenting the pertinent antecedents, this document presents the definition of the methodology and the techniques used, the results of the application of the techniques and, in conclusion, the estimation of the improvements offered by the system, once implemented.

The macro-objective of this work is to maximize the probability of monetization in the FC Barcelona app.

$$P(M) = \textit{Time Spent} * \textit{Frequency} * \textit{Retention}$$

The probability of monetization M is a product of the time that the user spends in the application (Time Spent), by the recurrence with which the user visits the application (Frequency), by the retention of said user in a timeline (Retention).

Background

As a starting point, we have reviewed the literature on standard methods and different approaches when dealing with recommendation problems [3] [4]. As a reference paper, we have taken the case of YouTube's recommendation system, considered a success story in the industry [2].

Project

The proposed system is based on three main entities: the user, the video and the interaction that involves them. A user registered in the digital ecosystem of F.C. Barcelona is defined by a series of demographic attributes (age, location, gender, etc.) and a set of attributes that describe its relationship with the ecosystem (date of registration, frequency of use, date of last use, type of user, etc.). Users can be non-registered, registered or paid. Non-registered users do not provide any information when using the App which offers little contribution and they have been removed from the data set for this work. Registered users have created an account and shared their demographic data, which is supplemented by the system with the activity log such as the first and last date of connection, frequency, duration of interactions, etc. These users are key since the main objective of the Club is to convert them into paying users. Paying users are users who, in addition to having a registration, make a monthly or annual payment and in return receive unrestricted, 24/7 access to BarçaTV+.

A video is defined by the type of its content (games, players, Club, etc.), the publication date, duration, access level (paid, open), sport, etc.

An interaction describes the viewing of a video by a user, also documenting the date and time of the video, the type of device, location, language, the time from and to the next game, and a number of other attributes.

An interaction is considered successful if the user viewed more than 30% of the length of the video, following industry standards [1].

In this work, the interactions of February 2021 involving registered and paying users were considered.

Figure 1 presents a schematic view of the structure of the methodology. Two main parts of it are evidenced, the analysis and creation of the knowledge base and the application of knowledge..

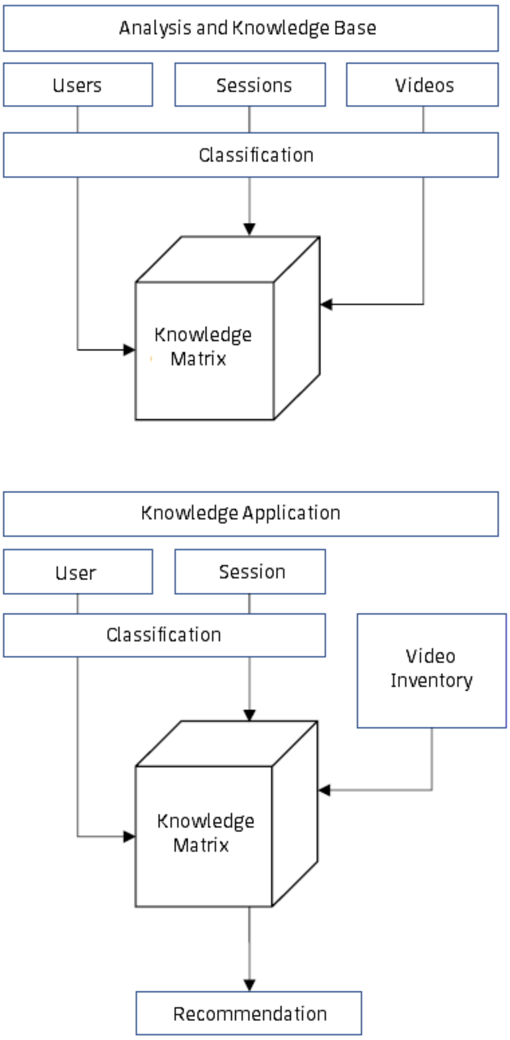

Figure 1 – Methodology Outline

The analysis and creation of the knowledge base consists of the unsupervised classification of the three entities (users, videos, interactions) from the observed combinations. The result is a cube of user, video, session combinations where it is also documented whether the majority of the times the combination was observed resulted in a successful display (more than 30% of the video was viewed) or not.

This allows answering questions such as: the times a class U user and a class S session were observed, what kind of video was successful most often?

The application of knowledge consists of the classification of users, videos and interactions in preparation for making a decision about what kind of video to show the user, given its class and the interaction class. The kind of video to show will be the one that has shown the best result according to the knowledge accumulated in the previous stage.

This will answer a question such as: if a user has been classified as U and her session as S, what kind of video should be presented to increase the probability of a successful viewing?

In a real implementation, users and videos would be constantly classified to be evaluated at the moment of decision making. It is important to note that, in both cases, temporary attributes such as the time since the user's last connection or the time since the video was published are constantly changing attributes and can classify the user or video into a different class as time passes. In other words, the classifications of users and videos are not permanent and must be carried out with a certain frequency, to be defined in the operational environment.

Sessions are by definition unique in time and they must be classified according to their attributes before being used for decision making.

The experimentation carried out during this project does not include the implementation of the system in an operational environment. Instead, an attempt is made to document the expected profit in terms of the percentage of videos viewed successfully, taking as a reference the day after the last observed in the data used for the creation of the knowledge base. The assumption for this is that the recommendation system would have been used as of March 1 based on the knowledge learned during February 2021.

Unlike the proposed three-dimensional methodology (sessions, users and videos), the experimentation showed that the influence of the session is negligible, making the problem two-dimensional. Hopefully this will change with a more detailed review of the attributes that sessions describe.

Data Processing

The experimentation was carried out with data extracted from information systems of

F.C Barcelona, once the necessary permits had been obtained.

The data set consisted of an interaction file and a user file. In both cases, the sets were reduced to contain only the rows relevant to the study.

The original file of interactions contained 445,474 rows from which those related to live videos and those with videos published more than 300 days ago (which are not planned to be promoted) were removed. This resulted in just over 380,000 conserved interactions.

From these, the videos and unique users were extracted to proceed with the classification. In total, 2,506 videos and 32,897 unique users were processed.

It is noteworthy that the same video can be considered multiple different instances as time passes, since, given the nature of the content, the interaction with them depends on the time elapsed since its original publication. For example, a summary of a soccer game is very interesting on the day of its publication, but it is less interesting two days later and practically irrelevant a week later. The 2,506 videos mentioned above consist of 613 videos that multiply over time and are re-viewed under their new age characteristics.

Once the three sets (instances, users and videos) were created, the attributes of each one were analyzed with the intention of reducing their dimensionality, ensuring that attributes were maintained that were not correlated with each other and that they had an influence on the characteristics of the each entity.

During the data processing of the interactions, it was concluded that the available data (time, application section, geographical location, type and size of the device, etc.) did not contribute relevant information to the fact that the interaction was successful or not. This led to the reduction of the model from three to two dimensions, with the user and the video being the relevant dimensions.

Attribute analysis techniques were used for users and videos, eliminating out-of-range and null values, separating categorical values into multiple “one-hot vectors”, keeping only one of several correlated attributes, converting all attributes to numeric and standardizing them to values. between 0 and 1.

Once these two data sets were constructed, the unsupervised K-means clustering method was used to identify classes from each set.

The combinations of user and video classes were then contrasted with the success values of each interaction in order to build a knowledge matrix.

Finally, a dataset was taken with the interactions from the beginning of March 2021 and the users and videos observed that day were extracted and classified. For each unsuccessful interaction, it was estimated what kind of video could be recommended to the user according to the previously constructed knowledge matrix, and from this the potential improvement in viewing time was estimated.

Experimentation

The experimentation process consisted of:

- Data extraction and preparation
- Attribute analysis
- Unsupervised classification (Clustering)
- Construction of the knowledge matrix
- Application of the model on data not used in training

During the extraction and preparation of the data, the user file and the interaction file were taken and from these the input data set for the classification processes was built.

The rows in each set that corresponded to the other were kept.

One-hot vectors were created for categorical variables such as user type, geographic location, language, and gender, among others. According to the number of categories, some of these variables were considered individually only in a maximum number, and the rest were grouped in "others". For example, the 3 languages with the most representation were taken and the rest were grouped into "otherLanguages". Users who have English, Spanish or Catalan as their registered language will have a value of "1" in their language and "0" in the other two and in "otherLanguages". Those who speak some other language will have "1" in "otherLanguages" and "0" in other languages.

Columns of dates were then rebuilt for conversion to a unique number format (Unix timestamp). Some were converted to elapsed time, such as age in years, rather than using date of birth.

Some attributes had no values and were filled in according to the case and need.

This process resulted in complete and numerical files for users, videos, and sessions, from which attributes were analyzed.

During attribute analysis, correlation matrices were used to eliminate attributes with high correlation with each other. Thus, out of 7 attributes initially identified for the videos, it was reduced to 4.

Figure 2 presents the correlation matrix for videos after cleaning the attributes. The first column is the identifier and the following are the age of the video in the system in days, duration, if it is open transmission and if it is paid. It should be noted that the negative scale does not reach -0.6, so even the dark cells indicate a low correlation between the variables.

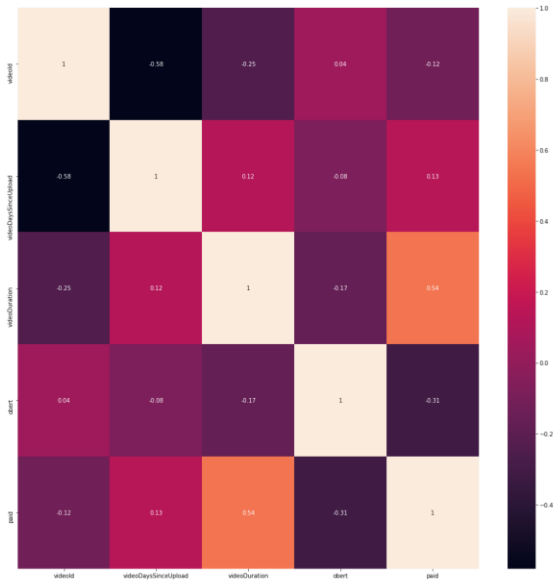

Figure 2 – Final Correlation Matrix (Videos)

In the case of the sessions, through the Correlation Matrix, the number of attributes was reduced only from 129 to 123, so it was also decided to apply the Mutual Information technique and compare the relevance of the attributes in terms of information. that contribute to the success of the visualization. The results yielded influence values in hundredths of a unit, and even in thousandths, leading to the conclusion that the characteristics of the sessions have no influence on the success of the visualization.

Figure 3 shows the result of the Mutual Information analysis of the 30 most influential variables, it can be seen the low value of the level of influence of the attributes on the success of the visualization.

This result led to the decision not to use the session properties as a dimension in the knowledge cube, reducing the model to two dimensions.

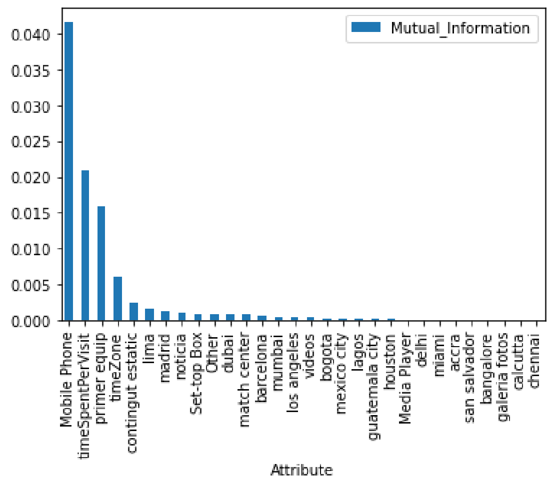

Figure 3 - Mutual information of session variables with visualization success

In the case of users, through the Correlation Matrix the number of attributes was reduced from 46 to 43.

Figure 4 presents the correlation matrix for users after attribute cleaning. The details are not expected to be read, it is presented to illustrate the complexity of the entity, based on the number of attributes.

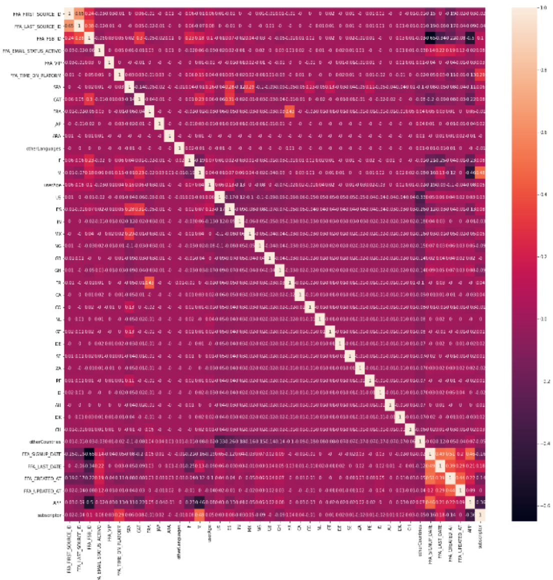

Figure 4 – Final Correlation Matrix (Users)

Once the data had been reduced due to its relevance, both in dimensions (only videos and users) and their attributes, the unsupervised classification was carried out.

The K-means clustering method was applied, executing with several K values until reaching a manageable value and such that higher K values did not offer visible improvements in the result.

The validation was done using the inertia method. *Figure 5* shows the inertia curve of the video clustering, where 8 was selected as the number of clusters.

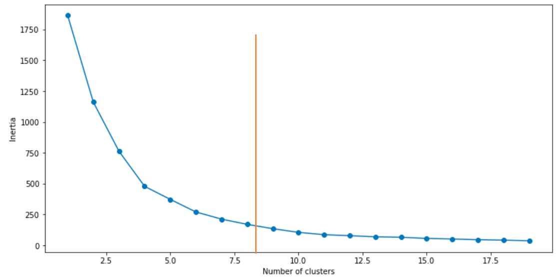

Figure 5 - Inertia elbow of video clustering

The result was a classification of 6 classes for videos and 8 for users.

With the results of the unsupervised classification, each interaction with the user and video class was labeled, thus building the Knowledge Matrix.

Table 1 shows the knowledge matrix for user class 0. It can be seen that the video class with the highest success rate is class 1 with 40% successful views. This knowledge suggests that when a class 0 user is in the application, a class 1 video should be recommended to them, if the chance of successful video viewing is to be maximized.

Video Class	User Class	Error	Success	Success Rate
0	0	700	73	9.44%
1	***0***	***2647***	***1798***	***40.45%***
2	0	838	186	18.16%
3	0	228	52	18.57%
4	0	458	215	31.95%
5	0	242	49	16.84%
6	0	1148	415	26.55%
7	0	39	18	31.58%

Table 2 – Knowledge matrix for class 0 user

It is estimated that a knowledge matrix similar to the one created will be valid for the interactions following the period analyzed. The extension in time of the period covered by the training data of the model would in itself be an object of future study.

In this work, the accumulated knowledge from interactions in February 2021 was used to estimate the potential improvement for the first days of March.

The users and videos that interacted with the application on March 1, 2021 were classified using the previously created models and the number of successes and failures was observed. Table 2 shows the summary of the interactions.

It can be seen, for example, that class 0 users viewed 44 class 1 videos resulting in 41 failures, a 6.82% success rate.

Video Class	User Class	Error	Success	Success Rate
0	0	8	4	33.33%
1	***0***	***41***	***3***	***6.82%***
2	0	8		0%
3	0	22	12	18.57%
4	0	2	3	35.29%
5	0	2	3	60.00%
6	0	2	0	0%
7	0	2	1	33.33%

Table 3 – March 2021 interactions

However, the conclusion in the knowledge matrix suggesting that showing class 1 videos to class 0 users would result in up to 40% successful viewing is negated in this March dataset.

Using a recommendation model based on the previously shown knowledge matrix, these users would have been recommended class 1 videos, resulting in a lower than expected success rate.

The reasons for this behavior may be related to the length of the period used for the generation of knowledge (one month), which may not be enough to model the behavior of users. This may have to do with the nature of video relevancy decline and possible high content variability that would affect video rankings. Added to this is the uncertainty described above regarding the fact that the characteristics of the session do not contribute to the model, a conclusion that is not aligned with business knowledge, especially in the correlation between days of play and user behavior.

The process was verified with other video and user classes and the behavior is similar.

Discussion

The methodology proposed in this work proved to be simple and repeatable, but not enough to make an accurate prediction. The lack of precision may be related to the volume of data used for model training, the variety of data, or the lack of more attributes in the original data set.

The classification process carried out was not supervised and may be improved as the methodology is applied and the process is refined from the construction of sets of videos and properly classified users.

During the attribute analysis, unexpected situations were discovered that must be studied in more detail, such as the fact that the characteristics of the session have no influence on the success of the video display. It would be expected that certain types of video have more or less relevance a certain number of hours before or after a game, or that certain types of videos result in a better visualization depending on the time of day and / or their content, but it is not known. found clear evidence of this.

Data problems were found that should be improved in an eventual automated implementation, such as data from incorrect dates, out of range or missing. The analysis also suggests that attributes that may have little value or are redundant among themselves (city, state, country) are being collected and maintained, while other more important ones, such as a unique encoding of the type of content of a video, are no longer used.

That said, the suggested methodology can serve as the basis for the development of a system of recommendations once the pending points open in the study have been considered.

Conclusion

The proposed recommendation methodology, based on the classification of users and videos, presents a potential but not yet proven way to improve the conversion of successful views.

The analysis carried out during the project questions the value of some attributes and highlights the need to capture others, or improve some already captured.

Starting from the initial unsupervised classification, work should be done on the definition and documentation of classes of videos and users that allow a formal and consistent classification.

The methodology proposes a video class to be recommended to a user according to its class at a given time. The specific video to be recommended (or the order in which several videos will be recommended) will result from an algorithm that combines the videos of that class available in the library with the business rules applicable at the time, considering promotion priorities of some types of content .

The study focused on the registered user and their conversion to a paid user. Converting from unregistered to registered users will require a different analysis, as there will be less data to classify and track users.

Learning from this work and the eventual implementation of the methodology should also lead to a detailed design of the content of the first 30% of each video, to improve the chances that it will be viewed for more than that time.

Bibliography

[1] Google Ad Manager Help (2021). *Overview of Viewability and Active View*. Retrieved from https://support.google.com/google-ads/answer/7029393?hl=en

[2] Paul Covington, Jay Adams, Emre Sargin (2016). 'Deep Neural Networks for YouTube Recommendations', *RecSys '16 September 15-19, 2016, Boston , MA, USA, 18-19 December.*

[3] Introduction to recommender systems (2019). *Overview of some major recommendation algorithms*. Retrieved from https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada

[4] Introduction to recommender systems (2020). *How services like Netflix, Amazon, and Youtube recommend items to the users?* https://towardsdatascience.com/introduction-to-recommender-systems-1-971bd274f421