**CCT College Dublin**

**Assessment Cover Page**

|  |  |
| --- | --- |
| **Program Title:** | M Sc Data Analytics |
| **Cohort:** | Part-time |
| **Module Title:** | Research and Professional Ethics |
| **Assessment Title:** | Integrated |
| **Lecturer Name:** | Christiani Eccher |
| **Student Full Name:** | Zygimantas Jakubauskas |
| **Student Number:** | Sba22342 |
| **Assessment Due Date:** | 05/May/2025 |
| **Date of Submission:** | 05/May/2025 |

**Declaration**

|  |
| --- |
| By submitting this assessment, I confirm that I have read the CCT policy on Academic Misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source. I declare it to be my own work and that all material from third parties has been appropriately referenced. I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution. |



**Evaluating Filtering Techniques in Recommender Systems for a Parent-Focused Activity App**

Author:

**[Zygimantas Jakubauskas]**

**[Sba22342]**

DATE: 05/05/2025

**Table of Content**

Table of Contents

[1. Introduction 4](#_Toc196068426)

[2. Literature Review 6](#_Toc196068427)

[3. Research structure 10](#_Toc196068428)

[4. Ethics and Legal Regulatory Considerations 13](#_Toc196068429)

[5. References 14](#_Toc196068430)

**List of Tables**

*Table1. Implementation plan outlined for getting the project done on time……………..6*

*Table2. Starting dataset at initial stage of the project…………………………………….14*

**List of Figures**

*Figure1. Research Methodology…………………………………………………………….13*

Abstract

*In this project the author analyses 5 recommendation filtering systems with focus on hybrid and knowledge-based recommendation filtering systems as less researched systems with the possible idea implementing the findings into the app helping parents to find suitable after-school activity for their children. Cons and pros discussed used latest research in this topic. Most common systems – collaborative and content-based analyzed extensively trying to identify ways of improving the results using alternative approaches.*

# Introduction

Recommender systems are now popular both commercially and in the research community, where many approaches have been suggested for providing recommendations. In the context of ML problems, recommender systems recommend data points to a consumer based on an inferred behavior pattern of that consumer (and the general behavior pattern of other consumers). Recommender systems are ubiquitous in the digital era – they exist everywhere, ranging from e-commerce websites to video sharing/streaming platforms, to social media. (Venkata Bhanu Prasad Tolety, 2022). These systems have been applied to many areas, such as movie recommendations, music recommendations, news recommendations, webpages and document recommendations. Many companies have employed and benefited from recommender systems, such as the book recommendation of [Amazon](https://www.sciencedirect.com/topics/earth-and-planetary-sciences/amazon), music recommendation of Apple Music, and product recommendation of TaoBao. (Donghui Wang, Yanchun Liang, Dong Xu , 2018).

With this in mind there is no reason why they can’t be used in finding activities for the kids. Keeping their children occupied is many parents’ worries. Finding suitable activities to keep them busy after school hours can be challenging. Imagine the family with kids moving into new to them area without any local knowledge. It takes time to make new friends for kids or parents to get that vital knowledge and insights. So, the app, in which parents could type in some relevant info in order to find most suitable solution for their family could potentially prove very useful.

* 1. **Research Motivation**

As all parents know there is a shortage of highly specialized apps helping not just to find clubs or other activities, but also camps, or local family-related festivals and other fun things to do. It would be great to have an app or a few apps, where your family’s needs could be stored, and recommendations would pop up as requested. Every parent can recall the moment when there is a weekend coming and there are no plans made simply because they have no information on what is happening in the city.

These kinds of apps could be created with the help of recommendation systems. These systems are improving all the time and when combined with AI other elements can really make many people’s lives easier. This project focuses on recommendation systems for outside-of-school activities for children but if findings are successful, they can be used for many other ideas (clubs for adults, summer camps for kids, local festivals).

* 1. **Research Objectives and/or Hypothesis**

There are number of *objectives* to be achieved in this research:

1. To evaluate different recommendation models, which work best to suit parents (and children) needs in the context of an activity app.
2. To evaluate the effectiveness of hybrid and knowledge-based filtering techniques in recommending suitable extracurricular activities.
3. Which approach can best target the very common problem known as “cold start”.
4. To explore the latest trends and discoveries in recommendation system development
5. To explore the opportunities to use the results for actual working app, that would help parents to find activities for kids in their local area outside school time. The creation of the app would be a bonus goal, as it falls a bit outside of data analytics domain and more in the programming area.

*Hypothesis*

1. Is knowledge-based model producing better results than a collaborative model.

It’s easier to succeed when there are clearly defined goals that are based in reality. These goals are described as SMART. [SMART in SMART goals stands for Specific, Measurable, Achievable, Relevant, and Time-Bound.](https://twitter.com/intent/tweet?source=webclient&amp;via=atlassian&amp;text=The%20SMART%20in%20SMART%20goals%20stands%20for%20Specific,%20Measurable,%20Achievable,%20Relevant,%20and%20Time-Bound.&amp;url=https://www.atlassian.com/blog/productivity/how-to-write-smart-goals)

Defining these parameters as they pertain to the goal helps ensure that objectives are attainable within a certain time frame. This approach eliminates generalities and guesswork, sets a clear timeline, and makes it easier to track progress and identify missed milestones. (Boogaard, 2023)

Is the goal **Specific**? In the Data Analytics domain, the goal is to analyse recommendation algorithms but most specifically knowledge-based algorithm.

Is the goal **Measurable**? 5 recommendation systems algorithms are selected for initial analysis, however collaborative and content-based are extensively researched by now and bigger focus will be given to newer models especially hybrid and knowledge-based.

Is the goal **Achievable**? All 5 algorithms already exist, no need to invent anything new. To analyse them and to determine what works best is an achievable goal.

Is the goal **Relevant**? Recommendation system algorithms are constantly evolving and improving. Even the biggest tech companies can still get them wrong. For example, Amazon can offer customer, who just bought the camera, to buy another camera. This can’t be happening, as customer already owns the camera, and most likely won’t need another one. Therefore, the problem is relevant.

Is the goal **Time-bound**? There is approximately 21 weeks allocated to complete the project from start to finish. If there is any amount of time saved during the period it can be used for analysis of more RecSys, improving the starting five using Deep Learning elements or creating the actual app.

A white table with black text

AI-generated content may be incorrect.

*Table1. Implementation plan outlined for getting the project done in time*

# Literature Review

* 1. **Introduction**

A recommender in system is a program that attempts to recommend the most suitable in terms to specific users by predicting a user’s interest. (J. Lu, 2015).

This project will be looking at 5 Recommender System (RecSys) models mainly, however there is a chance more can be explored along the way:

* Collaborative Filtering
* Content-based Filtering
* Knowledge-based Filtering
* Context-based Filtering
* Hybrid Filtering

**Collaborative Filtering**. In collaborative filtering (CF), users are modeled as a simple list containing the ratings provided by the user for some items. (Francesco Ricci, 2015). Ratings could be difficult to obtain for this project. Collaborative Filtering is the process of filtering or evaluating items using the opinions of other people. (S.Aarthi, 2013). Options are to use sentiment analysis using reviews from forums or social media platforms. Or they can be assigned by the author, but then there is a danger of bias. Another option is higher clubs in terms of member numbers get higher ratings.

**Content-based Filtering**. On the other end of the spectrum lie content-based filtering approaches. These approaches explicitly model the items being recommended and ensure that such items suit the profile of the users' preferences. These methods build a model of the user's preferences by examining the history of the user's interaction with the system, and an item-profile (a set of attributes or feature vectors computed for an item) (Venkata Bhanu Prasad Tolety, 2022)

**Knowledge-Based Recommendation System** (KBRS) distinguishes itself among the various types of RS by applying another technique to produce a recommendation. A KBRS generates recommendations based on the domain knowledge. A user will get a recommendation based on his profile and the behaviour of other users will not be considered at all, or when it is, it will not play a central role in determining the recommendation.

The KBRS can thus be used to address limitations of the common recommendation approaches. When using the knowledge-based approach, no large data set is necessary and the cold-start, new item and the grey sheep problem are thus avoided. Also, because the domain knowledge, on which are based the recommendations, is noise-free the recommendations are more reliable. The only limitation faced by the KBRS is the construction of the knowledge base, which usually is a complicated task that demands considerable domain knowledge, and expertise in knowledge representation. (Bouraga, 2014)

**Context Aware Recommendation Systems** take different [contextual attributes](https://www.sciencedirect.com/topics/computer-science/contextual-attribute) into consideration and try to capture user preferences correctly. (Saurabh Kulkarni, Sunil F. Rodd, 2020). In this project it is important to take into context the location.

**Hybrid Filtering**. Finally, one of the newer concepts is Hybrid filtering. By integrating content-based, collaborative, and knowledge-based techniques, the hybrid model notably enhances recommendation accuracy, effectively addressing challenges like the "cold start" problem and data sparsity issues. (Amangeldieva & Kharmyssov, 2024).

* 1. **Subtopics**

The two recommendation systems (Collaborative and Content-based) are to date extensively researched and most often used in today’s world. However, even when it comes to most popular algorithm – Collaborative – even to this day still there is no one single wide-spread method, which approach is working best. CF methods range from early matrix factorization to recently emerged deep learning based methods. (Xiang Wang, Xiangnan He, 2019)

However, there are few limitations for these techniques. Many authors highlight three major problems for collaborative filtering – cold start, sparsity, scalability.

The cold start problem is related to the lack of information (i.e., for users and items) available in the recommendation algorithm. Three types of cold start problems could be identified: (a) recommendations for new users, (b) recommendations for new items, and (c) recommendations on new items for new users. (Blerina Lika, Kostas Kolomvatsos, Stathes Hadjiefthymiades, 2014)

This problem will be faced with this project, as there will be many new users as well as new items (clubs). Collaborative authors, who presented their project in 2022 tried to solve new user problems using user ratings to predict user preference and they claim that it worked more effectively than traditional algorithms and popularity-based approach. A product is referred to as popular by the most user ratings in one specific product. (Farha Islam; Md Shohel Arman; Nusrat Jahan; Musabbir Hasan Sammak; Nusrat Tasnim; Imran Mahmud, 2022). However, sometimes items ratings may not be available too.

In [2016 IEEE International Conference on Big Data](https://ieeexplore.ieee.org/xpl/conhome/7818133/proceeding) 5 authors claimed they successfully tackled cold start problem using the novelty approach employing deep learning. One of the most important features of the proposed technique is the fact that it can be applied on top of any existing CF based recommendation engine without changing the CF core. They successfully applied this technique to overcome the item cold-start problem in CareerBuilder’s CF based recommendation engine. The experiments showed that the proposed technique is very efficient to resolve the cold-start problem while maintaining high accuracy of the CF recommendations. (Jianbo Yuan; Walid Shalaby; Mohammed Korayem; David Lin; Khalifeh AlJadda; Jiebo Luo, 2016). However, in newer research (2023) authors H. Yuan, and A. A. Hernandez analyzed 8 authors work using Deep Learning and came to conclusion that despite this method offers high accuracy it lack interpretability. Findings presented in table 14 of their report. (H. Yuan, A. A. Hernandez, 2023). Other authors add that a lot of data is needed to bootstrap recommendations, which is not available for a new item/new user. (Venkata Bhanu Prasad Tolety, 2022)

Sparsity is another problem and refers to the amount of available data or the density of data points in a dataset. In the context of recommendation systems, sparsity can refer to the number of ratings or interactions that users have had with items. If the ratings are very sparse in many applications, they cause CF-based methods to degrade significantly in their recommendation performance. (Hao Wang, 2015). The sparsity requirements for recommendation systems can vary depending on the type of system and the goals of the recommendations. (Camilleri, 2023)

So which RecSys approach would work best in the context of finding the best club to suit family needs. This is one of the objectives of this study.

Authors [Guy Shani](https://link.springer.com/chapter/10.1007/978-0-387-85820-3_8#auth-Guy-Shani) and [Asela Gunawardana](https://link.springer.com/chapter/10.1007/978-0-387-85820-3_8#auth-Asela-Gunawardana) claim that a first step towards selecting an appropriate algorithm is to decide which properties of the application to focus upon when making this choice. Indeed, recommendation systems have a variety of properties that may affect user experience, such as accuracy, robustness, scalability, and so forth. (Guy Shani, 2010)

The basic assumption in collaborative filtering is that people who agree at one point in time on a specific item will agree in the future on similar items. Moreover, it is assumed that people will continue to keep liking items similar in nature to what they liked in the past. (Venkata Bhanu Prasad Tolety, 2022)

Authors Xiang Wang,and Xiangnan He in 2019 also came to conclusion, that CF methods may be insufficient to provide good recommendations: Ranging from early matrix factorization to recently emerged deep learning based methods, existing efforts typically obtain a user's (or an item's) embedding by mapping from pre-existing features that describe the user (or the item), such as ID and attributes. We argue that an inherent drawback of such methods is that the collaborative signal, which is latent in user-item interactions, is not encoded in the embedding process. As such, the resultant embeddings may not be sufficient to capture the collaborative filtering effect. (Xiang Wang, Xiangnan He, 2019)

Conventional recommendation systems either use content based or collaborative filtering-based approaches to model user preferences and give recommendations. These systems usually fail to consider evolving user preferences in different contextual situations. (Saurabh Kulkarni, Sunil F. Rodd, 2020)

Here is when consideration of other methods comes to mind, first of all hybrid (combining best of few words) and knowledge-based (which may well be suited for this problem area but it still has to be yet to be proven).

In 2021 authors [Dhiraj Khurana](https://ieeexplore.ieee.org/author/37088878421) and [Sunita Dhingra](https://ieeexplore.ieee.org/author/37088878759) introduced improvement over the existing hybrid and knowledge based recommender system is proposed by integrating the clustering method within content based filter and classification method within collaborative filter. The proposed method handled the scalability problem by using the fuzzy clustering method. This reduced dimension-based dataset is processed by the probabilistic Bayesian network classifier for predicting the recommendations. The sparsity problem is handled in both stages of this model. The proposed recommender system model is applied on MovieLens dataset. The comparative analysis was done against content-based recommender system (CBRS), Pearson correlation based collaborative recommender system (PCRS), Frequency-weighted Pearson Correlation (FPC), Weighted Pearson Correlation (WPC) and hybrid recommender systems (HRS). The average RMSE rate achieved by CBRS, PCRS, FPC, WPC, HRS and the proposed hybrid recommender system are 0.3851, 0.3515, 0.3527, 0.3539, 0.3340 and 0.1987 respectively. The significant reduction in MAE rate is also identified in this work. The experimentation results identified that the proposed model reduced the error rate and improved the accuracy rate over existing systems. (Dhiraj Khurana; Sunita Dhingra, 2021)

* 1. **Conclusion**

There was an effort made to make literature sources as up to date as possible, trying to avoid anything written before 2010 as new ideas and approaches are constantly appearing but even though some claims “something will be used in the future”, are already being implemented. The topic isn’t overused or outdated, as authors cannot all agree on one single right method, and many approaches tried have their own advantages and disadvantages.

Three main tools were used for research purposes – Google Scholar, Research Rabbit and Zotero, all very helpful in terms of filtering literature, making it recent, quickly finding all the relevant information.

1. Research structure
   1. **Primary and Secondary Research**

There are many ways to collect primary data.

**Sentimental analysis**. Although considered one of the weaker because it doesn’t identify sarcasm, negation, grammar mistakes, misspellings, or irony (Dilmegani, 2025), it still can be helpful to especially in assigning ratings to different clubs or activities, based on parents reviews on different platforms. Ratings are vital for Collaborative filtering.

**Interviews, surveys**. As target audience are parents with school kids, interviews and surveys can be conducted to find out what factors influenced parents’ decisions to pick curtain clubs for their children, how many activities they are doing now, were there any wrong choices, when children lasted only few days or weeks in the club. Out of responses received following variables can be created: Parent’s ID (to keep data anonymized no real names will be used, only unique ID), Child’s ID (for same reasons as parent’s), Kid’s character (e.g. shy, active), Kid’s age, Character (e.g. shy, energetic), Area of interest ( e.g. Sport, Art, Outdoor), Locality (e.g. Artane, Castleknock), Radius (e.g. within 3km, within 5km), Happiness (how happy the child is at the current club, rated 1-5, with 5 representing Very happy), Previous Experience (e.g. 2 years Soccer), Choice (activity child is doing at the moment), Additional choice (additional activity child is doing). Realistically it is possible to collect data about few hundred kids. Variable Choice would become target variable.

**Simulated data**. Combined with the dataset (df\_interviews) of real data received from interviews and surveys, or forum observations, the rest of dataset would be randomly populated data (except of target variable). There are a few ways to input a target variable. It could be imputed by the author, although this would introduce bias, as this would be based on author’s educated guess, rather than real life event. Or dataset df\_interviews could serve as training dataset, and simulated dataset (df\_simulated) could serve as testing dataset. Classification machine learning models can be used then to predict target variable. As most of the independent variables are categorical, the best models to be used would be CatBoost, LightGBM, Decision Trees and Random Forest. CatBoost is especially efficient in predicting categorical features. (Ibrahim, 2020)

**Results**. Results calculated by the best performing machine learning model will be further used for RecSys algorithms.

As for secondary research, data of existing clubs can be easily obtained as it is publicly available from their own- websites or official social media accounts. The variables extracted would include Club (e.g. KUBS, Whitehall GAA), Activity (e.g. Irish Dance, Soccer, GAA, Taekwondo), Genre (e.g. Sport, Art, Martial Arts), Ages Catered (e.g. 9-16), Location (e.g. Artane, Castleknock), Gender (Boys, Girls, Both), Skill Level (All, Beginner, Advanced), Open (Weekends, Wednesday only).

So how much data is needed to get appropriate results?

Daniel Camilleri claims that there is no minimum amount of data required for a recommendation system; instead, what matters is the quality of interactions and contextual information about users and items being ranked. (Camilleri, 2023)

In one study authors claim using real dataset comprising of only 111 students organized into interdisciplinary groups. They claim the results showcase the clear benefits that our hybrid recommendation system enjoys, showing more than 30 percentage points of improvement over conventional filtering techniques. (Venkata Bhanu Prasad Tolety, 2022)

* 1. **Sampling Strategy**

The most obvious survey targets would be parents with active kids. But this restriction is not necessary here as the aim is not to find relationships between the activities say in period 2015-2025. The aim is to establish overall trends, how parents choose the activities never mind the era. Therefore, parents themselves can be interviewed if they were active themselves as children. Yes, it’s possible that in the 80’s some activities, like programming, computer classes or some newer martial arts classes didn’t exist, but if the number of people surveyed is quite big it shouldn’t affect the initial stage too much.

* 1. **Methods of Research**

Most interviews will be done in person, to ensure parents share the information willingly and to get more accurate information in general. Online surveying opportunities will be offered too but there is a danger that some people will not be suitable for this project, might skip questions, therefore creating a lot of missing data and in general large number of replies when a survey is done online, is not expected.

It is very important to capture parents from many walks of life, not just interview parents whose kids attend the very same club as that would introduce bias and incorrect information.

A diagram of data analysis

AI-generated content may be incorrect.

*Figure1. Research Methodology*

# Ethics and Legal Regulatory Considerations

* 1. **Legal/ Regulatory Considerations**

[Recommender systems](https://www.sciencedirect.com/topics/computer-science/recommender-systems) have developed in parallel with the web. They were initially based on demographic, content-based and collaborative filtering. Currently, these systems are incorporating [social information](https://www.sciencedirect.com/topics/computer-science/social-information). In the future, they will use implicit, local and personal information from the [Internet of things](https://www.sciencedirect.com/topics/computer-science/internet-of-things). (J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, 2013). This 4 authors work, written in 2013 is now 12-year-old, and social, local and personal information if attempted to be used today, may breach GPDR act which was only introduced in 2016.

When there is any talk involving children or data of children, there are always ethical considerations. However, it is important to note that this is not kids, but parents’ app. Parents themselves submit the info, they want to provide. If any of the parents doesn’t want to be involved, there is no problem with that Furthermore, no vulnerable information, e.g. real names or exact addresses, must be used. All research to be carried out will be completely anonymized.

However, just in case, is better to avoid wording such as “other parents in your locality also chose this club for their child” if there is a danger those parents or kids could be recognized. According to Children’s data and parental consent act from Data Protection Commission, this could be sometimes interpreted as sharing data, and sharing data without parents’ permission is forbidden except of certain situations. (Commission, 2023).

* 1. **Ethics and Data Protection**

If data is completely anonymized, then it is no longer personal data. Guidance on Anonymization and Pseudonymization issued by Data Protection Commission states that Irreversibly and effectively anonymized data is not “personal data” and the data protection principles do not have to be complied with in respect of such data. Pseudonymized data remains personal data. (Commision, 2019). Same act describes personal data. Personal data means any information relating to an identified or identifiable individual. This individual is also known as a ‘data subject’. An identifiable individual is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that individual. (Commision, 2019). Therefore, questions like name or exact address can’t and won’t be asked. To be on the safe side even names of the parents won’t be asked to avoid children being recognized, although for the project names aren’t needed anyway.

At the initial stage of the project dataset could look something like this:

A screenshot of a computer

AI-generated content may be incorrect.

*Table2. Starting dataset at initial stage of the project*

As can be seen from the snippet provided, it’s impossible to recognize any child or even parent involved in the project. At the initial stage of the project not even location isn’t necessary as first step is to recommend a type of activity, not the actual club.

Another thing to be considered is the use of clubs’ data for the application. Even though this data is available and easily accessible from their official websites or social media accounts, clubs may not necessarily want it to appear on the app, especially commercial app. Collecting general club data, like location or opening hours is legal, but collecting sensitive info, like email is not and is described in GPDR act. Furthermore, rules on how to use or share clubs’ data can often be found in relevant clubs’ websites and especially if the data is used commercially, clubs permission must be obtained unless stated otherwise on club’s website or official social media account.

# References

Amangeldieva, A. & Kharmyssov, C., 2024. *A hybrid approach for a movie recommender system using Content-Based, Collaborative and Knowledge-Based Filtering methods.* Astana, IEEE.

Blerina Lika, Kostas Kolomvatsos, Stathes Hadjiefthymiades, 2014. Facing the cold start problem in recommender systems. *Expert Systems with Applications,* 41(4), pp. 2065-2073.

Boogaard, K., 2023. *https://www.atlassian.com.* [Tinkle]   
Available at: https://www.atlassian.com/blog/productivity/how-to-write-smart-goals#:~:text=The%20SMART%20in%20SMART%20goals,M:%20Measurable  
[Kreiptasi 21 04 2025].

Bouraga, S., 2014. Knowledge-Based Recommendation Systems: A Survey. *: International Journal of Intelligent Information Technologies,* 10(2), p. 19.

Camilleri, D., 2023. *https://www.shaped.ai.* [Tinkle]   
Available at: https://www.shaped.ai/blog/part-2-how-much-data-do-i-need-for-a-recommendation-system#:~:text=We%20highlighted%20that%20there%20is,users%20and%20items%20being%20ranked.  
[Kreiptasi 26 04 2025].

Commision, D. P., 2019. *https://www.dataprotection.ie.* [Tinkle]   
Available at: https://www.dataprotection.ie/sites/default/files/uploads/2019-06/190614%20Anonymisation%20and%20Pseudonymisation.pdf

Commission, D. P., 2023. *https://www.dataprotection.ie.* [Tinkle]   
Available at: https://www.dataprotection.ie/sites/default/files/uploads/2023-04/DPC\_ChildrensData\_ParentalConsent.pdf  
[Kreiptasi 20 04 2025].

Dhiraj Khurana; Sunita Dhingra, 2021. *An Improved Hybrid and Knowledge Based Recommender System for Accurate Prediction of Movies.* New Delhi, India, IEEE.

Dilmegani, C., 2025. *https://research.aimultiple.com.* [Tinkle]   
Available at: https://research.aimultiple.com/sentiment-analysis-methods/  
[Kreiptasi 20 04 2025].

Donghui Wang, Yanchun Liang, Dong Xu , 2018. A content-based recommender system for computer science publications. *Knowledge-Based Systems,* 157 tomas, pp. 1-9.

Farha Islam; Md Shohel Arman; Nusrat Jahan; Musabbir Hasan Sammak; Nusrat Tasnim; Imran Mahmud, 2022. *Model and Popularity Based Recommendation System- A Collaborative Filtering Approach.* Kharagpur, India, IEEE.

Francesco Ricci, L. R. a. B. S., 2015. *Recommender Systems.* 2 mont. New York: Springer.

Guy Shani, A. G., 2010. *Evaluating Recommendation Systems.* 2 mont. s.l.:Springer.

H. Yuan, A. A. Hernandez, 2023. : User Cold Start Problem in Recommendation Systems: A Systematic Review. *IEEE Access,* 11 tomas, pp. 136958 - 136977.

Hao Wang, N. W. D.-Y. Y., 2015. *KDD '15: The 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* Sydney, Association for Computing Machinery.

Ibrahim, A. A., 2020. Comparison of the CatBoost Classifier with other. *(IJACSA) International Journal of Advanced Computer Science and Applications,* 11(11), p. 96.

J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, 2013. Recommender systems survey. *Knowledge-Based Systems,* 46 tomas, pp. 109-132.

J. Lu, D. W. M. M. W. W. a. G. Z., 2015. A survey,” Decision Support Systems". *Recommender system application developments,* 74 tomas, pp. 12-32.

Jianbo Yuan; Walid Shalaby; Mohammed Korayem; David Lin; Khalifeh AlJadda; Jiebo Luo, 2016. *Solving cold-start problem in large-scale recommendation engines: A deep learning approach.* Washington, IEEE.

S.Aarthi, R. a., 2013. Collaborative Filtering Approach Based Recommender Systems. *International Journal of Engineering Research & Technology (IJERT),* 2(8), p. 4.

Saurabh Kulkarni, Sunil F. Rodd, 2020. Context Aware Recommendation Systems: A review of the state of the art techniques. *Computer Science Review,* 37 tomas.

Venkata Bhanu Prasad Tolety, E. V. P., 2022. Hybrid content and collaborative filtering based. *Bulletin of Electrical Engineering and Informatics*, 06, pp. 1543-1549.

Xiang Wang, Xiangnan He, 2019. *Neural Graph Collaborative Filtering.* Paris, Association for Computing MachineryNew YorkNYUnited States.

**GitHub link:** https://github.com/ZygimantasJakubauskas/Research-and-Ethics