

A Recommendation Mechanism for a Non-Fungible Token Trading Platform

Pai-Ching Tseng^{*}, Teck-Xu Cheng[†], Yu-Jing Jang[†],
Tzu-Ching Weng[†], Iuon-Chang Lin[†]

Abstract

Technology has evolved at an increasing pace over the decades, blockchain is used in many fields, and Non-Fungible Token (NFT) is one of its applications. With NFT, all digital assets can be freely traded because NFT establishes a unique token for each asset. From the transaction statistics, it can be found that the transaction volume in the past six months has increased explosively, and the projects uploaded to the NFT trading platform have also increased; but it also means customers need to spend more time finding what they want. Active recommendation systems can be used to solve this problem to reduce search cost. This study proposes a recommendation system for NFT tokens, compares the similarity between items, and analyzes buyers' collection data to prove that the recommendation theory proposed in this study is valid.

Keywords: Non-Fungible Token, Recommendation system, Cosine Similarity

1 Introduction

In recent years, the blockchain (BC) has developed rapidly. The initial application is the Bitcoin cryptocurrency. Traditional transactions require a third-party organization to record all information, and blockchain technology is essentially a "decentralized" database, which is composed of blocks and chains. Each block records itself and the hash value of the previous block. Comparing these hash values can verify whether the data has been changed. Blockchain technology has been applied to many fields, such as cryptocurrency [1], smart contracts [2], Internet of Things [3].

Ethereum is similar to Bitcoin as a cryptocurrency platform currency platform, both are based on blockchain technology. Ethereum introduced smart contract technology, so various applications can be built on the platform and various smart contracts can be formulated in Ethereum to ensure the security of transactions [2][4]. NFT is a cryptocurrency for smart contracts. Each NFT is bound to the digital data on the platform, such as painting, music, video, etc, and establishes its own smart contract. Users can trade and transfer through NFTs.

Both NFT and Bitcoin are cryptocurrencies, but the values of their respective tokens are different. Bitcoin is the equal value, but NFT represents different values due to the different binding projects. In the past, digital works were difficult to trade because the Internet made data easy to copy and circulate, and it was impossible to prove transaction records and data ownership after purchase. Now the above problems can be solved through the proof of smart

^{*} Ph.D. Program of Business, Feng Chia University, Taichung, Taiwan

[†] Department of Management Information Systems, National Chung Hsing University, Taichung, Taiwan

contracts and a trusted trading platform. In the future, the NFT market can have a better development [5][6].

With the popularization of NFTs, the types of transactions are becoming more and more diverse, such as Movies NFTs, Artworks NTFs, Internet of Things NFTs [7]. NFT trading platforms have also developed such as Cryptoslam, Opensea, Nifty Gateway, Zora. According to Dune Analytics statistics, using the Opensea platform for statistics, since August 2021, the monthly transaction volume is 2 billion to 3 billion US dollars as shown in Figure 1, with nearly 200,000 monthly users as shown in Figure 2 to trade on the platform [8]. Although the NFT market is a new market, from the transaction volume and the number of transactions, it can be seen that when the society gradually accepts this transaction model in the future, there will be more users and more types of NFTs entering the market.

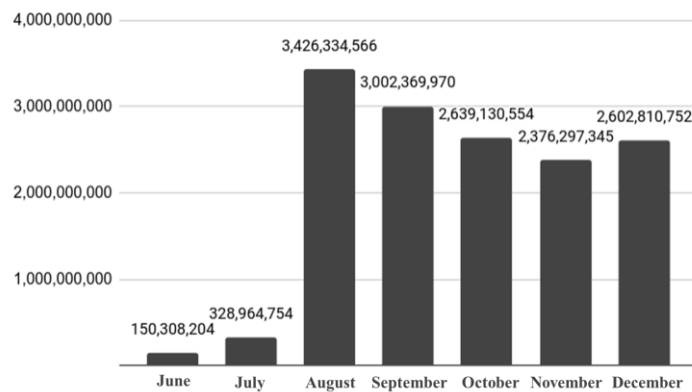


Figure 1: Opensea Transaction Volume 2021 (USD)

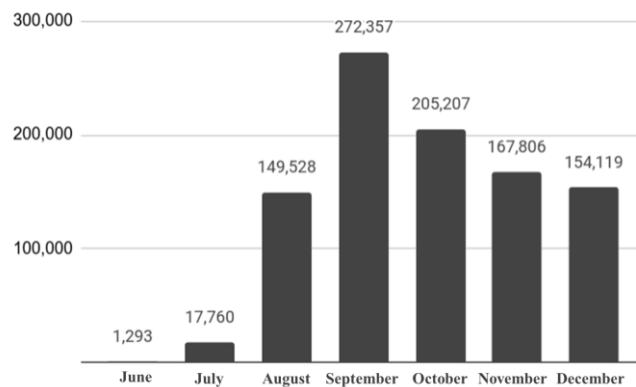


Figure 2: The number of Opensea users who have participated in at least one transaction in the current month

In the era of internet information explosion, when the number of pending transactions increases, users need to spend more time searching for their favorite products. How the system can help buyers to purchase products more easily, increase transactions, and reduce search costs, is a problem that many scholars are concerned about. Therefore, some scholars use a recommendation system to solve this problem [9].

When the internet becomes popularized, people are accustomed to searching for information on the Internet. A large volume of information circulates on the internet, but users need to spend a lot of time searching for information. For example, E-commerce [10],

streaming video platforms [11], social platforms [12], learning video recommendations and other websites with a large volume of information all face the same problem, so they launched the recommendation system and achieved good results. The system actively recommends content to users to improve user search efficiency [13].

Opensea is an NFT trading platform with a large volume of transactions recently. Taking the Opensea web page as an example, when users use the NFT platform, they use tags to search for information, such as creator, category, category details, volume, and etc. The platform displays the results that match the search condition, and the user checks from the search results if there is a favorite item.

2 Related Works

2.1 Non-Fungible Token

Non-Fungible Token (NFT) is a cryptocurrency extended by smart contract. A unique identification certificate is established by binding the token to the digital data. Both NFT and Bitcoin are cryptocurrencies, but each has a different value. Bitcoin is the same value, but NFT represents different values due to the different binding projects. According to the research results of Dowling [14], NFT and cryptocurrencies are low correlation asset types. In the past, digital works were difficult to trade because the internet made data easy to copy and circulate, and it was impossible to prove transaction records and data ownership after purchase. Therefore, digital assets are less likely to be collected than physical assets. When NFTs became popular, creators could bind their works to NFTs, and they could use tokens to prove the existence of digital works, as well as their ownership. The value of the token is affected by a number of factors such as uniqueness, creator, year of creation, rarity, etc. Ante [15] states that NFTs create opportunities for digital data with unique value.

2.2 Recommendation System

In recent years, with the vigorous development of the Internet, various websites have been established one after another, and the transmission and retrieval of information has also become easier. When a large volume of data is uploaded to the internet, users need to spend more time searching for data, and it is more difficult to find what they want. Although the internet has a lot of convenience to modern people, the huge volume of data invisibly increases the search cost, that is Information Overloading; this is a common topic of discussion [16]. Recommendation system is the solution to this problem. Recommendation systems can help users overcome information overload on the Internet and help users make decisions [17]. Ricci [18] believes that e-commerce websites can use recommendation systems to suggest products or provide information to help consumers effectively make shopping decisions. The information filtering mechanism proposed in order to reduce the extra cost of the user in the process of searching for information and help users make decisions. The recommendation system can recommend potential information, services or products that the user may need according to the user's preferences, interests or behaviors [19], in order to improve the user's decision-making performance and quality, as well as to obtain higher customer satisfaction and loyalty. Schafer et al. [20] believed that

recommendation systems can guide users to make more effective purchasing decisions by recording the knowledge of users' preferred products.

Common recommendation system methods are as follows:

- a. Content-based Filtering is to collect user historical data, analyze past purchase records, operation records, viewing records, questionnaires, ratings, etc and analyze users' interests, preferences or past favorites, purchase records and other information to set user data. The advantage of this recommendation is that it can be recommended according to the user's preferences, but the disadvantage is that only product attributes can be considered. When various attribute characteristics are the same, the recommendation order cannot be determined. Lops et al. [21] pointed out that Content-based Filtering only considers the user's own past preferences, the recommendation scope can only be limited to the highly relevant products that the user has purchased in the past, and the selection of item attributes is limited.
- b. Collaborative Filtering analyzes the user's rating data, and categorizes users with similar preferences into the same group to find out the preferences of other users in the group as a basis for recommendation [22]. Scholars such as Herlocker et al. [23] also mentioned that collaborative filtering connects a group of people with the same interests, and predicts recommendations through group data.
- c. Hybrid Recommendation Systems; it is a combination of "Content Based Filtering" and "Collaborative Filtering" that combine the two filtering methods to address their respective shortcomings and turn the advantages of each method into a recommendation system. Scholars such as Geetha et al. [24] used a hybrid recommendation system for movies, and compared with the filtering algorithm alone, the hybrid recommendation system accuracy was higher.

3 Research Methods

This study selects the Opensea NFT trading platform for research discussions and analyzes the top 20 creators by trading volume (data extracted in December 2021). There are many types of works, including animals, characters, creatures, codes, sandbox maps, houses, etc. The over specialization of the recommendation system is one of the most frequently discussed issues. When user preferences and product categories are too unique, Collaborative Filtering classification is not very useful, because similar groups cannot be classified, resulting in inaccurate recommendations. After analyzing the top 20 creators in this study, it is believed that NFTs such as code and houses are too unique. This study hopes that the differences between NFTs are not too large and the research field is too divergent, so they are not included as training data. Those NFTs with similar characteristics are characters, creatures, etc. Therefore, this study has set four inclusion conditions, namely characters, animals, and creatures with facial features. There are more than 500 trading works, which limit the field of this research data.

In this research, 200 pieces of data are randomly selected from each NFT collected as training data, and a total of 1,600 pieces of data are obtained. Every piece of work created by the creator is different. After binding the NFT, a unique existence is created. Although

NFTs are different, under the constant creation of creators, the works are similar in style. Opensea's NFT transaction structure is as follows. After creating digital data, creators create NFTs through the platform system, and create corresponding features for each NFT. Therefore, the system will provide "NFT feature search function" to increase users' search efficiency. After the search is complete, a list of matching NFTs will be listed, which the user can then choose.

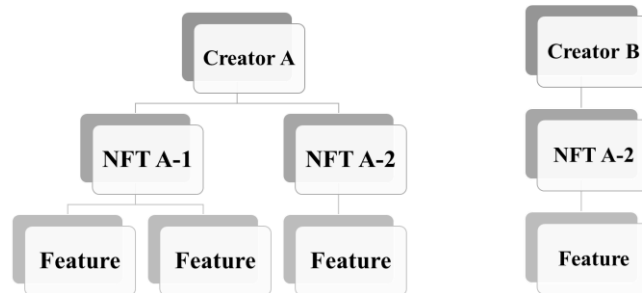


Figure 3: Opensea creator, NFT, feature relationship architecture diagram

Now, NFT platforms do not provide recommendation systems, so this research will try to find ways to recommend them. NFT is most suitable for collaborative filtering analysis, by comparing features, using cosine similarity algorithms to calculate similarity and statistical similarity.

$$\text{Cosine Similarity} = |a \cap b| / \sqrt{|a| * |b|}$$

This research analyzes the collected NFT items through the purchase records of users, and compares the similarity between them. When the user has two similar NFTs, it proves that the recommendation hypothesis proposed in this study holds, and the system can recommend similar information to the user.

4 Research Implementation Analysis and Design

Research data collection; In this study, 8 creators were selected as training dataset, and 200 materials were randomly obtained from each, a total of 1,600 materials were obtained. Use the classification field of NFT as a training feature, and provide the detailed content of the obtained product through Opensea API. Use Http Curl to call the "retrieving a single asset" api to obtain the detailed information of the product. Opensea responds data in JSON format for users to use. In this study, categorical data will be intercepted, so the two fields "trait_type" and "value" are obtained from the "traits" field. The classification data is created by the creator, and the amount of data is unlimited.

Data structuring; After analysis, the proper noun thesaurus is extracted, and the text data is converted into structured data. When the text data is to be structured, use the proper noun dictionary established by this study to search for the root, and find all the roots with the same beginning. Suppose the first word of the article is "Black", there are three root words "Black", "Black eye", and "Black long hair" that match the conditions. This study uses the "longest radical method" to select the radicals. Select the matching radical with the highest number of words. Only if the above condition is met, the root can be completely replaced.

Recommendation system establishment; Collaborative Filtering gathers everyone's comments, opinions, and ideas, and analyzes the similarity between users and items through statistical analysis, and then serves as the basis for recommendation. Each NFT represents a digital work, and each work is an independent existence, so this study will use the item characteristics of NFT for similarity comparison. There are three steps in this stage: establishing the feature set of the recommendation system, training the recommendation system model, and analyzing the user's collection.

Build a recommendation system feature set; In this phase, we will do data preprocessing, data structuring, and binding of products and functions. The recommendation system compares the features of the items, and calculates the similarity between the items through the algorithm. The product features of the NFT are the classification features established by the creators themselves. Therefore, classification name and classification data are used as the training feature of the recommendation system in this study. After observing the newly added features, this study only analyzes the similarity of words, thus excluding symbols, such as ";", "[number]", "(,)", "-", "", ".", etc.

Recommendation system model training; This study decided to study the similarity of items and used collaborative filtering to analyze the similarity between items. First, extract and sort out the features of the NFT to be trained Table 1. The first step is to invert the feature with the NFT Table 2. The second step is to count the number of intersections between each NFT Table 3.

Table 1: NFT feature correspondence table

	Black Eye	Female	Male	Black Hair	Number Of Features
NFT A	1	1	0	1	3
NFT B	1	1	0	0	2
NFT C	0	0	1	1	2

Table 2: Features, NFT comparison table

Feature	NFT
Black Eye	NFT A 、 NFT B
Female	NFT A 、 NFT B
Male	NFT C
Black Hair	NFT A 、 NFT C

Analyzing the user's collection; There is also the owner's information in the NFT detailed information obtained through Opensea API. The user's information on the NFT platform is hidden and represented by encrypted codes. There are 1600 training data in this study, so 200 NFT owners were randomly obtained, and 50 data were randomly obtained from their collections for verification, and the owners with less than 5 collections were removed from the verification test data Table 4.

Table 3: Statistics of the number of NFT intersections

	NFT A	NFT B	NFT C
NFT A	x	2	1
NFT B	2	x	0
NFT C	1	0	x

Table 4: Owner's corresponding data sheet

Owner			Owner NFT		
id	user_token		id	owner_id	nft_token
1	USER_A	↔	1	1	NFT_A
2	USER_B		2	1	NFT_B
			3	1	NFT_C
			4	2	NFT_D

In order to verify whether the owner will purchase similar NFTs, it is necessary to cross-count all NFTs with the most similar NFTs, sort the statistical results, and record the highest and lowest similarities.

5 Research Results

User Collection Similarity Analysis; In this study, 200 users were randomly selected for analysis. By collecting their NFT data and the similarity between different NFTs, calculate the highest slope of the two items in the NFTs collected by each user. The statistical results are shown in Figure 4 below. The statistical chart shows that users will buy products with high similarity.

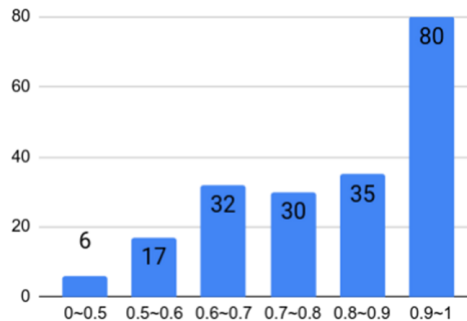


Figure 4: Similarity statistical graph of the number of user collections

Recommendation Analysis Statistics; From Table 5 below, it can be found that 70% of the data in the “full data search model” are recommended by the system with a data similarity greater than 0.7, and only a few data cannot find similar data. The recommendation rate of

"different creator model" is opposite to that of model one. After removing the condition of the same creator, only 10% of the data can find a similarity greater than 0.7.

Table 5: Recommendation system analysis slope quantity statistics table

The Slope Of The Recommended Product	Model 1 Full Data Model Result	Model 2 Different Creator Model Result
< 0.5	6	1,014
≥ 0.5	1,594	586
≥ 0.6	1,545	444
≥ 0.7	1,168	184
≥ 0.8	436	17
≥ 0.9	61	4

Full data search model; Full data search model obtains the training results, and extracts the least similar (slope < 0.5) data and the most similar (slope ≥ 0.9) data for comparison.

The least similar data makes up less than 1% of the total training data. After observation, this research believes that it should be a new series of products launched by the creator, and there is no similar data, which will result in a low recommendation rate. Through the recommendation results, this study found that the statistical slope of 70% of the recommended ratios was greater than 0.7, and the slope of the distribution of 90% of the recommended ratios was greater than 0.6. The same creators are included in the full data search recommender model, so a large percentage of the same creators are recommended.

Different creator model; Different creator model obtains training results, and extracts the least similar (slope < 0.5) data and the most similar (slope ≥ 0.9) data for comparison. The least similar data accounts for more than 60% of the total training data, and this recommendation result is within the expectation of this study. Although the similarity is very low, some features are similar, so it can lead to other creators. The most similar data accounts for less than 1% of the total training data. Although it is NFT released by different creators, it can be found that the recommendations of this group are the same series of data. The recommended result with a slope average of 0.7 is the same series.

Integrated analysis; The integrated analysis is to take out the product to be recommended, take out the full data search model and the different creator model, and compare the two recommendation results. After analysis, the following two points are found:

- Full data search model: Many are from the same creator, because their characteristics will be similar.
- Different creator model: Usually there are similar feature classification names and will be judged to be similar.

Recommendation system accuracy; Through the trained recommendation system, this study counted users who collected multiple items at the same time through 1600 training

data, and obtained 54 users as verification data. The number of transactions is shown in Table 6 below. The number of collections for each user is divided into training set and validation set, and three cutting ratios are used for training, 8020 (80% training set, 20% validation set), 7030 (70% training set, 30% validation set), 6040 (60% training set, 40% validation set). Through the recommendation system of this study, five types of data (10, 20, 30, 40, and 50) are recommended to each training set at the same time, and the recommended data is compared with the validation set data. The number of successful recommendations is used to calculate the accuracy recommended by this study.

Table 6: Statistical table of the number of user transactions

Number Of User Purchases	Number Of Users
24	1
22	1
15	1
14	1
12	1
11	3
10	1
9	2
8	3
7	1
6	1
5	3
4	11
3	24
Total	54

The statistical results are as follows: Table 7, Table 8, Table 9, 8020 (80% training set, 20% validation set statistics table), 7030 (training set 70%, validation set 30% statistics table), 6040 (60% of the training set and 40% of the validation set statistical table), each training set is recommended by the system to 10, 20, 30, 40, 50 items, and the number of accuracy is counted.

Table 7: Statistics of 80% training set and 20% validation set

	10 recommended	20 recommended	30 recommended	40 recommended	50 recommended
0	35	28	25	22	19
0~0.6	6	8	5	2	2
0.6~0.7	2	1	3	2	2
0.7~0.8	0	0	0	0	0
0.8~0.9	0	1	0	1	0
0.9~1	11	16	21	27	31

Table 8: Statistics of 70% training set and 30% validation set

	10 recommended	20 recommended	30 recommended	40 recommended	50 recommended
0	34	26	21	17	15
0~0.6	13	11	10	9	6
0.6~0.7	0	3	3	2	4
0.7~0.8	0	1	3	3	3
0.8~0.9	0	0	1	2	1
0.9~1	7	13	16	21	25

Table 9: Statistics of 60% training set and 40% validation set

	10 recommended	20 recommended	30 recommended	40 recommended	50 recommended
0	36	29	24	20	17
0~0.6	17	18	18	21	20
0.6~0.7	1	5	8	2	1
0.7~0.8	0	1	0	1	3
0.8~0.9	0	0	1	5	5
0.9~1	0	1	3	5	8

The recommended accuracy ratios are shown in Table 10. It can be seen from Table 10 that the accuracy of the 80% training set combination is lower than that of the 70% training set combination because the validation dataset and training dataset are too small. And 70% of the combinations in the training set have the best recommendation ratio. After recommending 30 to users, the accuracy of successful purchases exceeds 60%, so it is the best combination.

Table 10: Recommended Accuracy Table

	10 recommended	20 recommended	30 recommended	40 recommended	50 recommended
80% Training 20% Validation	35.19%	48.15%	53.70%	59.26%	64.81%
70% Training 30% Validation	37.04%	51.85%	61.11%	68.52%	72.22%
60% Training 40% Validation	33.33%	46.30%	55.56%	62.96%	68.52%

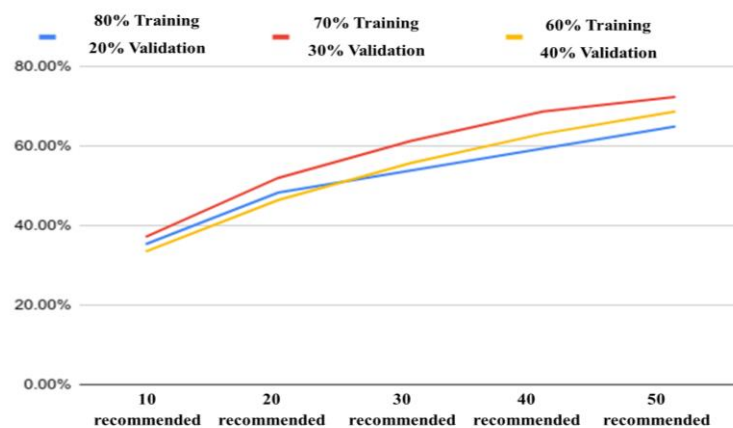


Figure 5: Recommended accuracy comparison chart

6 Conclusion

This study discusses the recommendation system of NFT, and establishes two recommendation models, which are "full data search model" and "different creator model". The following conclusions are drawn from the experimental results:

- By analyzing the purchase records of Opensea users, it is found that most users will purchase similar products. This result is also the motivation of this study and can prove the recommended hypothesis of the study.
- The recommendation results of the full data search model, the proportion of recommending similar is higher than that of the different creator model. The most similar recommendation comes from the same creator, and the least similar is also the same creator. This research believes that it should be a newly established NFT, and there is no similar creation yet.
- Recommendation results from different creators, the recommended ratio is very low, but the items recommended by observation still have corresponding similar characteristics, so they can be used to recommend other creators.
- From the accuracy of the recommendation system, it can be found that using the recommendation system in this study, 30 items are actively recommended to users, and 60% of users can be successfully purchased, so it proves that the cosine similarity recommendation method proposed in this study is valid.

References

- [1] Nakamoto, Satoshi. "Bitcoin: A peer-to-peer electronic cash system." *Decentralized Business Review* (2008): 21260.
- [2] Wood, Gavin. "Ethereum: A secure decentralised generalised transaction ledger." *Ethereum project yellow paper 151.2014* (2014): 1-32.
- [3] Huh, Seyoung, Sangrae Cho, and Soohyung Kim. "Managing IoT devices using blockchain platform." *2017 19th international conference on advanced communication technology (ICACT)*. IEEE, 2017
- [4] ANTE, Lennart. "Smart contracts on the blockchain—A bibliometric analysis and review." *Telematics and Informatics*, 2021, 57: 101519.
- [5] WANG, Qin, et al. "Non-fungible token (NFT): Overview, evaluation, opportunities and challenges". *arXiv preprint arXiv:2105.07447*, 2021.
- [6] FAIRFIELD, Joshua AT. "Tokenized: The law of non-fungible tokens and unique digital property". *Ind. LJ*, 2022, 97: 1261.
- [7] SGHAIER OMAR, A.; BASIR, Otman. "Capability-based non-fungible tokens approach for a decentralized AAA framework in IoT". In: *Blockchain Cybersecurity, Trust and Privacy*. Springer, Cham, 2020. p. 7-31.
- [8] Dune Analytics, Artificial Intelligence Blog; <https://dune.xyz/home>.
- [9] LU, Jie, et al. "Recommender system application developments: a survey". *Decision Support Systems*, 2015, 74: 12-32.
- [10] LINDEN, Greg; SMITH, Brent; YORK, Jeremy. "Amazon. com recommendations: Item-to-item collaborative filtering". *IEEE Internet computing*, 2003, 7.1: 76-80.

- [11] PONNAM, Lakshmi Tharun, et al. "Movie recommender system using item based collaborative filtering technique". In: 2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS). IEEE, 2016. p. 1-5.
- [12] CHEN, Rui, et al. "A survey of collaborative filtering-based recommender systems: From traditional methods to hybrid methods based on social networks". IEEE Access, 2018, 6: 64301-64320.
- [13] SHARMA, Lalita; GERA, Anju. "A survey of recommendation system: Research challenges". International Journal of Engineering Trends and Technology (IJETT), 2013, 4.5: 1989-1992.
- [14] DOWLING, Michael. "Is non-fungible token pricing driven by cryptocurrencies?". Finance Research Letters, 2022, 44: 102097.
- [15] ANTE, Lennart. "The non-fungible token (NFT) market and its relationship with Bitcoin and Ethereum". FinTech, 2022, 1.3: 216-224.
- [16] EDMUNDS, Angela; MORRIS, Anne. "The problem of information overload in business organisations: a review of the literature". International journal of information management, 2000, 20.1: 17-28.
- [17] HÄUBL, Gerald; TRIFTS, Valerie. "Consumer decision making in online shopping environments: The effects of interactive decision aids". Marketing science, 2000, 19.1: 4-21.
- [18] RICCI, Francesco, ROKACH, Lior SHAPIRA, Bracha. "Introduction to recommender systems handbook". In: Recommender systems handbook. Springer, Boston, MA, 2011. p. 1-35.
- [19] RASHID, Al Mamunur, et al. "Getting to know you: learning new user preferences in recommender systems". In: Proceedings of the 7th international conference on Intelligent user interfaces. 2002. p. 127-134.
- [20] SCHAFER, J. Ben, et al. "Collaborative filtering recommender systems". In: The adaptive web. Springer, Berlin, Heidelberg, 2007. p. 291-324.
- [21] LOPS, Pasquale, et al. "Trends in content-based recommendation". User Modeling and User-Adapted Interaction, 2019, 29.2: 239-249.
- [22] SARWAR, Badrul, et al. "Item-based collaborative filtering recommendation algorithms". In: Proceedings of the 10th international conference on World Wide Web. 2001. p. 285-295.
- [23] HERLOCKER, Jonathan L.; KONSTAN, Joseph A.; RIEDL, John. "Explaining collaborative filtering recommendations". In: Proceedings of the 2000 ACM conference on Computer supported cooperative work. 2000. p. 241-250.
- [24] GEETHA, G., et al. "A hybrid approach using collaborative filtering and content based filtering for recommender system". In: Journal of Physics: Conference Series. IOP Publishing, 2018. p. 012101.