Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.33.35.45

Personalized Travel Recommendation Systems: A Study of Machine Learning Approaches in Tourism

Mohamed Badouch[1], Mehdi Boutaounte[2*]

[1]*Faculty of Sciences, Ibn Zohr University, Agadir, Morocco*
[2*]*National School of Commerce and Management, Ibn Zohr University, Dakhla, Morocco*

Email: [1]mohamed.badouch.ext@edu.uiz.ac.ma,
Corresponding Email: [2]m.boutaounte@uiz.ac.ma*

Received: 29 December 2022 **Accepted:** 25 March 2023 **Published:** 26 April 2023

***Abstract:** Recommender systems that utilize machine learning algorithms are a prominent tool in the design and implementation of personalized tourism experiences. These systems analyze user data to generate recommendations for destinations, attractions, accommodations, and activities based on user preferences, behavior, and similarity to other users. Collaborative filtering and content-based filtering are two widely used machine learning algorithms in recommender systems, and hybrid systems that combine both approaches have shown to be effective in producing more accurate recommendations. Tourism recommendation systems (TRS) provide several benefits, including personalization, convenience, improved user experience, and increased revenue for tourism businesses. These systems can suggest destinations, attractions, accommodations, and activities that match user preferences and past behaviors, ultimately simplifying the trip planning process. Machine learning algorithms can be trained on large datasets to generate personalized recommendations, and can continuously improve their effectiveness by incorporating new data and user feedback. This paper provides a state-of-the-art overview of various types of recommendation systems (RS), including those based on user preferences, behaviors, demographic profiles, and social network judgments. The paper also presents a comparison table for these approaches. Additionally, the paper discusses the different stages of the travel process and the sources of data that can be used to develop a recommender system. The concluding section of the paper highlights the importance of personalized recommendations in the tourism industry and the potential for future research in this area.*

***Keywords:** Machine Learning, Recommender Systems, Tourism.*

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.33.35.45

1. INTRODUCTION

Machine learning is one of the key techniques used in recommender system design and implementation. A recommender system is a tool that recommends items to users based on their preferences, behavior, or similarity to other users. Machine learning algorithms are used to analyze patterns and relationships in the data to make these recommendations.
There are several types of machine learning algorithms commonly used in recommender systems, including:

- Collaborative filtering: This is an algorithm based on the similarity between users and their previous interactions with items.
- Content-based filtering: This is a recommendation algorithm based on article attributes and user preferences.
- Hybrid recommender systems: This is a combination of collaborative filtering and content- based filtering, often resulting in more accurate recommendations.

Machine learning algorithms are trained on large data sets to generate recommendations, and can improve the effectiveness of recommendations over time by updating the algorithms with new data and user feedback.
A TRS can be used to make personalized suggestions for travel destinations, attractions, accommodations, and activities based on a user's preferences and past behaviors. The goal of a travel RS is to help users plan their trips by making informed suggestions that match their interests and needs.

There are many possibilities that machine learning can be used to build a RS for tourism:
- Attraction Recommendation: Attractions, such as museums, historical sites, and natural parks, based on a user's interests, such as history, art, or nature.
- Destination Recommendation: Travel destinations based on a user's preferred climate, budget, and type of activity, such as sightseeing, adventure sports, or relaxation.
- Accommodation Recommendation: Accommodations, such as hotels, resorts, or vacation rentals, based on a user's budget, preferred amenities, and location.
- Activity Recommendation: Activities, such as tours, events, and experiences, based on a user's interests, budget, and location.
- These recommendations can be made using a com- bination of collaborative filtering, content-based filtering, and deep learning techniques. The RS can also take into account external factors, such as seasonality, weather, and travel restrictions, to make more benefits. There are some examples of benefits:
- Personalization: RS provide personalized suggestions for travel destinations, attractions, accommodations, and activities based on a user's preferences and past behaviors. This makes it easier for users to plan their trips and find destinations and activities that match their interests and needs.
- Convenience: RS make it easier for users to find the information they need to plan their trips. By providing personalized suggestions, users don't have to spend as much time searching for information or making decisions.

- Improved User Experience: RS can improve the overall user experience by providing relevant and personalized suggestions. This can help users find the information they need more quickly and easily, leading to a better overall experience.
- Increased Revenue: By providing relevant and
- personalized suggestions, RS can increase revenue for tourism businesses. For example, a RS that suggests accommodations and activities can increase bookings and sales for hotels, resorts, and tour operators.

This paper presents at first a state of the art on the various types of RSs, not only based on preferences of the users, but also on the basis of their behaviors, their demographic profiles, and the judgments of other users (presented on the social networks), a table that provides a comparison of these approaches is presented. In the third paragraph, we will focus on the various stages of travel and sources of data, and then the stages of processing them in a RS, we conclude in Section 4 on the conclusion of our work

2. Recommender Systems, Comparison of Types

2.1 Content-based recommendation

A Content-based recommendation is a widely used approach for personalized recommendations in which items are recommended to users based on their preferences for certain features or content. This approach operates under the assumption that users have similar preferences to the items they have previously liked or interacted with [1].

In a content-based RS, items are described using a set of features or attributes, such as genre, director, or actor, and recommendations are generated based on the similarity between these features and the user's past preferences. This method has been widely used in various domains, and has been shown to be effective in providing personalized recommendations to users. One of the advantages of content-based recommendation is that it does not require explicit feedback from users, as it only needs to know what items they have interacted with, making it a scalable solution for large datasets.

However, the accuracy of content-based recommendations can be limited by the quality and availability of item features, and by the potential for users to change their preferences over time [2].

2.2 Collaborative filtering based recommendation

Collaborative filtering-based recommendation is a widely-used approach in the field of RSs. It aims to provide personalized recommendations to users based on their past behaviors and the behaviors of other users.

In collaborative filtering, the system generates recommendations by identifying similar users or items and making predictions based on these similarities [1]. The two main types of collaborative filtering are user-based and item-based. User-based collaborative filtering generates recommendations based on the similarities between users, while item-based collaborative filtering generates recommendations based on the similarities between items. Both approaches have been shown to be effective in generating accurate recommendations, and they have been widely adopted in a variety of applications, including music and movie recommendations, travel recommendations, and e-commerce recommendations.

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.33.35.45

Despite the benefits of collaborative filtering-based RSs, they also have some limitations, such as the cold start problem, the sparsity issue, and the scalability problem.

2.3 Hybrid recommender system

A hybrid recommender system combines multiple recommendation algorithms to provide a more personalized and accurate recommendation to users [3]. The use of multiple algorithms can overcome the limitations of a single algorithm, such as over-reliance on a single type of information or the inability to handle multiple types of data. By integrating various algorithms and data sources, a hybrid recommender system can provide a more comprehensive and nuanced recommendation to users, leading to a better user experience and improved satisfaction.

Additionally, the use of multiple algorithms can provide a more robust recommendation, as the system can compensate for any biases or limitations in a single algorithm [4]. Overall, we believe that hybrid recommender systems have great potential to revolutionize the field of travel recommendation [4].

Table. 1 represents the different advantages and disadvantages of each system.

Approach	Advantages	Disadvantages
Content-based	Can make recommendations for users with unique or well-defined preferences. Can handle a diverse range of items.	Can suffer from the ’cold start’ problem, where it can be difficult to make recommendations for new users with limited data. Can lead to limited diversity in recommendations, as the system may only recommend items with similar attributes to those the user has liked in the past.
Collaborative filtering	Can handle a large number of users and items. Can provide recommendations for users with more general preferences	Can suffer from the ’sparsity’ problem, where it may be difficult to find similar users for certain items or users. Can be influenced by the behavior of a small number of influential users.
Hybrid filtering	Can handle a diverse range of items and a large number of users. Can overcome the limitations of single-algorithm systems. Can provide a more nuanced recommendation, taking into account both the user’s personal preferences and the preferences of similar users.	Can be more complex to design and implement compared to single-algorithm systems. Can be less interpretable, as the system’s decision-making process may be harder to understand.

Table 1: Table of comparison

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.33.35.45

2.4 Conclusion
Content-based, collaborative filtering and hybrid RS are three common techniques for generating personalized recommendations. Content-based RS rely on attributes of proposed items, such as: Recommendations are made on features. This approach uses users' previous preferences and interests to make suggestions. On the other hand, collaborative filtering RS make suggestion based on the behavior and preferences of similar users. This approach exploits the similarity between users to provide suggestion that are more likely to be widely accepted. Hybrid type combine the advantages of content-based and collaborative filtering systems, using attributes of items and behaviors of similar users to generate decisions.
These systems can overcome the limitations of either approach by combining the advantages of both methods. For example, hybrid RS can provide accurate propositions even when a user's preferences are not well-defined by using the preferences of similar users as a fallback.

3. Recommender system in Tourism

3.1 Stages of travel
Before travel:
Any applications of machine learning in tourism are related to the pre-travel phase, such as: Accommodation (localization, cost, confirmation…), Flight (price, luggage…), Travel planning and so forth [5]. Travel planning is an essential task for anyone who wants to explore new places or take a break from everyday life. With the advent of recommender systems, travel planning becomes more efficient and personalized. A popular approach to building TRS is to use algorithms such as Markov models, discrete hidden Markov model [6] or neural networks. Markov models are probabilistic models that can be used to analyze sequences of events and predict future events based on past events. In the context of travel planning, Markov models can be used to predict the most likely sequence of destinations based on a traveler's past travel patterns. Neural networks [7], on the other hand, are a class of machine learning algorithms that can be used to learn patterns in large datasets.
In the context of travel planning, neural networks can be trained on large datasets of travel itineraries to learn patterns and provide travelers with personalized recommendations based on their preferences and past behavior [5]. Overall, using algorithms such as Markov models and neural networks can significantly improve the accuracy and efficiency of itinerary planning RS

During travel:
Perhaps the most used stage of RS is during travel, as this stage carries variables and requirements for the user, whether during his wandering to new places to visit.
With the rise of smart cities and the ubiquity of mobile devices, RS and decision support systems have become powerful tools for enhancing the travel experience. By leveraging user data and machine learning algorithms, RS can suggest personalized travel itineraries, restaurants, attractions, and activities based on users' preferences, location, and past behavior. In addition, Decision Support Systems (DSS) can help travelers make more informed choices by providing real-time information on traffic, weather, events, and other relevant factors. By combining these two systems, travelers can optimize their travel plans, save time, reduce stress, and discover new experiences that they may have otherwise missed.

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.33.35.45

For instance, a traveler can use a RS to suggest a museum to visit then check traffic and find the most efficient route to get there. As smart cities continue to evolve, RS will play an increasingly important role in helping travelers navigate the urban environment and enjoy a more personalized and rewarding travel experience

After travel:
After Trip tourist reviews can be analyzed by various machine learning methods and useful information can be extracted from them.
By analyzing tourists' reviews and feedback, RS can learn about their preferences and interests, and suggest suitable travel options, such as destinations, accommodations, activities, and restaurants [8].
For instance, TripAdvisor's RS uses machine learning algorithms to analyze millions of user reviews and ratings to provide personalized recommendations to users based on their preferences [9]. Similarly, Booking.com's Genius program uses ML to analyze customers' booking history, preferences, and behaviors to offer them personalized discounts and perks. Furthermore, Airbnb's Price Tips feature uses ML to suggest optimal prices for hosts based on the supply and demand in their area.
Overall, RS and ML techniques have the potential to revolutionize the travel industry by providing tourists with more personalized and relevant travel recommendations.

3.2 Data sources
Users (text, image…):
Data sources like user text, images, videos, and other multimedia content generated by users can be utilized to design RSs. User text can provide valuable information about a user's preferences, opinions, and interests, which can be used to generate personalized recommendations.
For example, text data from user reviews can be analyzed using natural language processing techniques to identify patterns and preferences, which can then be used to recommend destinations, accommodations, and activities. Similarly, image and video data can be analyzed using computer vision techniques to identify visual features like landscapes [10], land- marks, and scenery, which can be used to recommend similar destinations or activities. For instance, a user who has shown a preference for beach destinations in their past travels might receive recommendations for other coastal locations based on the visual similarities between their past vacation photos and other potential destinations.

Gadgets and devices (position GPS, mobile roaming, Wifi. . .) :
With the rapid development of smart cities, it has become crucial to use data sources like gadgets and devices to gather information about user preferences and behavior.
For example, a tourist visiting a new city could use a mobile app that integrates GPS data to suggest nearby landmarks or restaurants based on their location and previous activity [8, 9]. Similarly, data from mobile roaming can provide information about the user's past travel history, which can be used to make personalized recommendations based on their previous experiences. By analyzing these data sources, it is possible to create a tailored TRS that takes into account the unique preferences and interests of each individual user. For instance, a TRS

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.33.35.45

could suggest hiking trails for a user who frequently visits mountainous areas, or recommend restaurants with vegan options for a user who has a history of ordering vegetarian meals.

Operations (web search, online booking…):
Incorporating operations such as web search and online booking into such systems can enhance their recommendation capabilities and provide more personalized recommendations to tourists. Web search operations can be used to gather in- formation about tourists' preferences, such as their favorite destinations, activities, and accommodation preferences. For instance, search engines and social media platforms can be used to collect data about tourists' behavior and preferences. Online booking data can also be used to improve tourism RS [7]. For example, online booking platforms can be used to recommend hotels, flights, and other services to tourists. By analyzing online booking data, we can learn which hotels are popular among tourists and which are not.
To illustrate the use of these operations in a TRS, consider the following example: A tourist wishes to plan a trip to New York City and provides their preferences to the RS. The system uses web search operations to gather information about the tourist's preferences and online booking data to recommend hotels, flights, and activities that fit their budget and preferences.

3.3 Phases of data processing (Example: Collaborative filtering - Hotel Recommender System)

Table 2: Dateset.

User id	Hotel id	Rating
1	291	2.0
2	98	5.0
3	162	1.0
1	635	1.0
5	725	4.0
6	8	3.0
3	61	2.0
8	97	4.0

Pre Processing:
To begin our analysis of hotels ratings, we must first transform the raw data into a "utility matrix". This matrix will represent each user's ratings of each hotel as a numerical value, allowing us to easily perform matrix-based operations and analyze the underlying patterns of user preferences. Therefore, the transformation of the hotel rating dataframe into a utility matrix is a critical step in our analysis, and one that must be executed with precision and care to ensure the validity and reliability of our subsequent analyses.

<table>
<tr><th colspan="8">Items</th></tr>
<tr><td rowspan="3">Users</td><td></td><td>1.0</td><td></td><td></td><td>2.0</td><td></td><td></td></tr>
<tr><td></td><td>3.0</td><td></td><td>5.0</td><td></td><td>4.0</td><td></td></tr>
<tr><td></td><td></td><td>1.0</td><td></td><td></td><td></td><td></td></tr>
</table>

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.33.35.45

	2.0						
							4.0

Table 3: User-item (utility matrix).

Model Training:
Once the data has been pre-processed, the next step in building a RS is to construct a model that can predict users' preferences. Matrix factorization is a widely-used technique within collaborative filtering [11]. However, there are other methods as well, such as neighborhood methods, that can be used to construct models. In the following sections, we outline the steps involved in using matrix factorization for collaborative filtering in a RS.
In order to model user-item interactions in RSs, it is common practice to factorize the user-item matrix into two latent factor matrices - a user-factor matrix and an item-factor matrix [11]. This approach al- lows us to capture the underlying latent factors that drive user-item preferences and to represent them in a lower-dimensional space. By decomposing the original matrix into two lower-rank matrices, we can reduce the dimensionality of the problem and obtain more meaningful representations of the user and item features.
Ratings are generated by humans and serve as features that are directly observable and assumed to be important. However, there are also a set of features that are not directly observable but still important in predicting ratings. These hidden features are called latent features and play a critical role in improving the accuracy of hotel RSs. Latent advantages may include aspects such as the hotel's architecture, location, and proximity to necessary facilities…

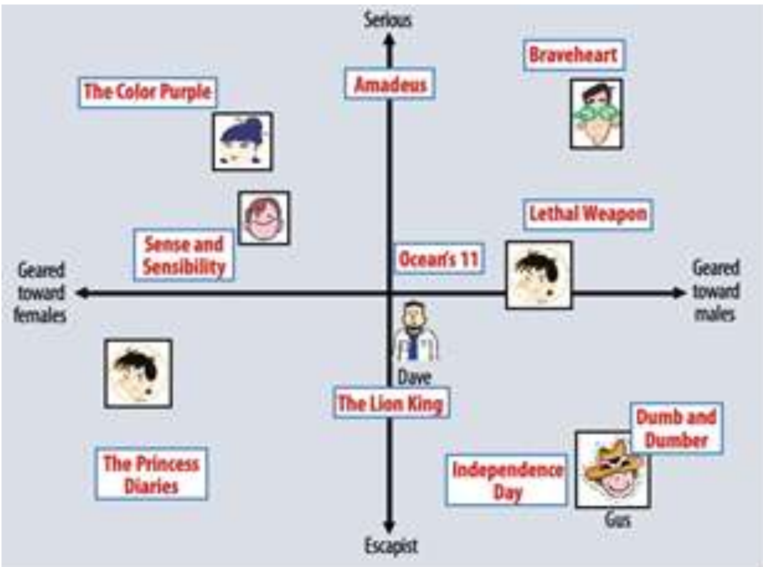

Figure 1: Latent factor approach-finds connections between users and objects to get new associations.., https://www.researchgate.net/figure/Latent-factor-approach-finds-connections-between-users-and-objects-to-get-newf ig2331552273

The concept of Latent Features [Figure 1] has emerged as a powerful tool in the field of RSs. Latent Features are essentially the underlying features that explain the interactions between

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.33.35.45

users and items in a RS. These features are not explicitly known to us, but can be inferred through statistical models such as matrix factorization.

Missing scores are predicted from the inner product of these two latent matrices [Figure 2]. Latent factors are typically represented by the variable K. By reconstructing the original user-item matrix using K, we can fill in the missing values and obtain predicted ratings for items that the user has not yet rated. Thus, the reconstructed matrix effectively populates the empty cells in the original user-item matrix, pro- viding a complete set of ratings and enabling more accurate recommendations to be made.

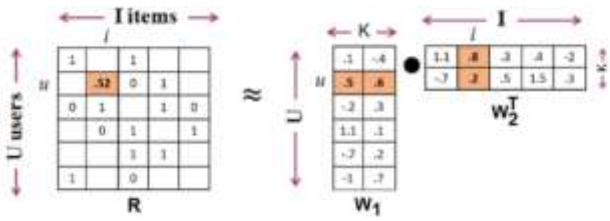

Figure 2: An example of matrix factorization, https://www.researchgate.net/figure/An-example-of-matrix-factorizationf ig1314071424.

There are various ways to implement the Matrix Factorisation using one of the following methods [12]:

- Alternating Least Squares (ALS)
- Stochastic Gradient Descent (SGD)
- Singular Value Decomposition (SVD)

Empty cells actually represent new users and new hotels. Therefore, if the proportion of new users is high, we may reconsider using some other recommendation methods, such as content-based filtering or hybrid filtering.

Optimization:

To effectively tune the parameters of a recommender system, it is important to select an appropriate evaluation metric. Precision@K [13, 14] is a commonly used evaluation metric that measures the proportion of relevant recommendations in the top K results provided to a user. The goal is to identify the optimal set of parameters that maximizes the Precision@K or any other relevant evaluation metric. Once the optimal parameters have been determined, model can be retrained to generate more accurate predicted ratings, which can be used to generate more effective recommendations for users.

Post Processing:

After predicting the ratings of items that a user has not interacted with, we can sort them in descending order and select the top N recommendations for the user. However, to avoid recommending items that the user has already interacted with, we need to exclude or filter them out. In the case of hotels, it is not worth- while to recommend a hotel that the user has already booked or expressed dislike for in the past [15]. This filtering step is essential for ensuring the relevance and usefulness of the recommendations and can im- prove the overall user experience. RS incorporate a filtering mechanism that takes into account the user's previous interactions and excludes those items from the recommendation list.

Evaluation:
Evaluating recommender systems is a crucial aspect of the development process. Conducting tests in real-world settings using techniques like A/B testing [16] is one of the most effective ways to evaluate these systems, as they provide feedback from real users. However, this may not always be feasible. In such cases, offline evaluation methods are employed.
We use a technique where we randomly mask some known ratings in the matrix. The masked ratings are then predicted using machine learning, and the predicted rating is compared with the actual rating. This allows us to evaluate the model's performance on the unseen data and estimate how well it can generalize to new users and items.

4. CONCLUSION

Recommender systems remains a very vague and forked field, it continues to develop with the emergence of smart cities and IoT. In this article, we discussed various recommendations types, while noting, for each approach, the pros and cons of using it, and based on that, the best approach can be chosen. In this paper we paid attention especially to RS in tourism, with specifying travel details and sources of information used during the stages of the process.
To delve deeper into RS, we made an overview example to show the line of data processing, filtering and extraction of suggestions using collaborative filtering with an indication of the obstacles that may be encountered.

5. REFERENCES

1. Khalid Al Fararni et al. "Hybrid recommender system for tourism based on big data and AI: A conceptual framework". In: Big Data Mining and Analytics 4 (Mar. 2021), pp. 47–55.
2. Umair Javed et al. "A Review of Content-Based and Context-Based Recommendation Systems". In: International Journal of Emerging Technologies in Learning 16 (2021), pp. 274–306.
3. Bankim Patel Verma Jai Prakash and Atul Pa- tel. "Big data analysis: recommendation system with Hadoop framework." In: IEEE Inter- national Conference on Computational Intelligence Communication Technology (2015), pp. 92–97.
4. Khalid Al Fararni et al. "Comparative Study on Approaches of Recommendation Systems". In: Advances in Intelligent Systems and Computing 1076 (2020), pp. 753–764.
5. Fatemehalsadat Afsahhosseini and Yaseen Al- Mulla. "Machine Learning in Tourism". In: ACM International Conference Proceeding Series (Sept. 2020), pp. 53–57.
6. Khatereh Jahromi and V.G. Jahromi. "Using Discrete Hidden Markov Model for modelling and forecasting the tourism demand in Isfahan". In: Journal of Information Systems and Telecommunication 6 (Mar. 2018), pp. 112–118.
7. Leyla Gamidullaeva et al. "A Design Concept for a Tourism Recommender System for Regional Development". In: Algorithms 16 (Jan. 2023).

Journal of Artificial Intelligence, Machine Learning and Neural Network
ISSN: 2799-1172
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JAIMLNN
DOI: https://doi.org/10.55529/jaimlnn.33.35.45

8. Joan Borràs, Antonio Moreno, and Aida Valls. “Intelligent tourism recommender systems: A survey”. In: Expert Systems with Applications 41 (Nov. 2014), pp. 7370–7389.
9. Alan Menk, Laura Sebastia, and Rebeca Ferreira. “Recommendation Systems for Tourism Based on Social Networks: A Survey”. In: arXiv preprint arXiv:1903.12099 (Mar. 2019).
10. William C. Gartner. “Image formation process”. In: Journal of Travel and Tourism Marketing 2 (Feb. 1994), pp. 191–216.
11. Hoang Nguyen-Tan, Hung Huynh-Huu, and Hiep Huynh-Xuan. “Collaborative filtering recommendation in the implication field”. In: International Journal of Machine Learning and Computing 8 (June 2018), pp. 214–222.
12. Gess Fathan, Teguh Bharata Adji, and Ridi Ferdiana. “Impact of matrix factorization and regularization hyperparameter on a recommender system for movies”. In: International Conference on Electrical Engineering, Computer Science and Informatics (EECSI) 2018-October (Oct. 2018), pp. 113–116.
13. Shaowei Wang et al. “EnTagRec ++: An enhanced tag recommendation system for soft-ware information sites”. In: Empirical Software Engineering 23 (2 Apr. 2018), pp. 800–832.
14. Jiawei Chen et al. “Fast Adaptively Weighted Matrix Factorization for Recommendation with Implicit Feedback”. In: Proceedings of the AAAI Conference on Artificial Intelligence 34 (04 Apr. 2020), pp. 3470–3477.
15. Mohd Abdul Hameed et al. “Collaborative Filtering Based Recommendation System: A sur- vey”. In: Article in International Journal on Computer Science and Engineering (2012).
16. “Offline A/B testing for recommender systems”. In: WSDM 2018 - Proceedings of the 11th ACM International Conference on Web Search and Data Mining 2018-Febuary (Feb. 2018), pp. 198– 206.