IRJCS: International Research Journal of Computer Science
Volume 11, Issue 02, February 2024
ISSN: 2393-9842
https://www.irjcs.com/archives

A Data-Infused Approach to optimal Paying Guest / Hostel Selections

Deepak Mangal
Department of Information Technology
Greater Noida Institute of Technology (Engineering Institute), India
Dr.Ajay Kumar Sahu
Department of Information Technology
Greater Noida Institute of Technology (Engineering Institute), India
ajay4989@gmail.com
Dr.Shivani Dubey
Department of Information Technology
Greater Noida Institute of Technology (Engineering Institute), India
dubey.shivani@gmail.com
Prof.Vikas Singhal
Department of Information Technology
Greater Noida Institute of Technology (Engineering Institute), India
Vikassinghal75@gmail.com
Rohit Ram Sandeep Singh
Department of Information Technology
Greater Noida Institute of Technology (Engineering Institute), India
Naim Akhtar
Department of Information Technology
Greater Noida Institute of Technology (Engineering Institute), India

Publication History
Manuscript Reference No: IRJCS/RS/Vol.11/Issue02/FBCS10085
Research Article | Open Access | Double-Blind Peer-Reviewed | Article ID: IRJCS/RS/Vol.11/Issue02/FBCS10085
Received:31, January 2024| Revised: 06, February 2024 | Accepted: 011 February 2024 | Published Online: 29, February 2024
http://www.irjcs.com/volumes/Vol11/iss-02/04.FBCS10085.pdf
ArticleCitation:Deepak,Ajay,Shivani,Vikas,Rohit,Naim(2024).A Data-Infused Approach to Optimal Paying Guest / Hostel Selections. IRJCS: International Research Journal of Computer Science, Volume 11, Issue 02 of 2024 pages 92-98.
https://doi.org/10.26562/irjcs.2024.v1102.04
BibTeX Deepak@2024Data
Academic Editor-Chief: Dr.A.Arul Lawrence Selvakumar,AM Publications, India

Copyright: ©2024 This is an open access article distributed under the terms of the Creative Commons Attribution License; Which Permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract: The world's population is expanding quickly every day, and a great number of individuals are moving to big cities and other locations in search of employment, better living conditions, or educational opportunities. The development of rental websites is covered in this paper. Property details can be uploaded by landlords and tenants can rent properties through this website. In this research, we have developed website will assist people in giving or accepting rental properties without the need to interact with apartment brokers in order to ultimately come to a mutually beneficial agreement. Among the one hundred different options, the project's goal is to identify and suggest the PG that a user is most likely to book.

I. INTRODUCTION

The global epidemic has had an impact on nearly every industry, with the hotel sector being particularly heavily struck. A suitable location to live has been significantly impacted, but the pandemic has also altered how we work, learn, and communicate. Everybody wants to live in a clean, safe environment, yet because of the pandemic, there are fewer homes that require proper maintenance. One of the first things to do when studying abroad from home is to make reservations for a decent location to stay. Choosing among thousands of properties when booking a vacation rental online can be a daunting undertaking. Motivated by the significance of these circumstances, we decided to use our expertise in the duty of PG recommendation to students [1]. Students who live far from their families for work will benefit from this endeavour as well. Two categories of people will be part of our target audience: individuals seeking rental housing and landlords with properties offered for rent. Students and working professionals can choose the place they are interested in after securing the desired location. The goal of this project is to create an interactive website for PG booking, specifically for students. It will also have a recommendation system for user convenience and an interface that allows property owners to add or remove properties. The concept behind the project is that everyone can obtain PG from our website at any time they need it [2]. Each and every detail regarding the PG, including its characteristics and price, is available on our website. It also includes information about the rooms' availability:

a) Recommendation Systems, Current Research

Trends, and Upcoming Initiatives In recent years, recommendation systems have become a prominent field of study with the goal of helping users discover recommendations that match their interests on the internet.

This research provides a regressive analysis of the recommendation system, spanning several information gathering strategies and suggestion approaches. A large number of academics from all over the world are interested in researching it because of its many uses. This paper's primary goal is to identify current trends in the recommendation field. From 2011 to 2016, around 1000 research papers were published and examined by various publications. This study produced a number of fascinating results that will aid current and future academics in their analysis and selection of research directions [3].

b) System of Recommendations for Scholarly Writing

The degree of the recommendation system has increased as a result of the research paper's vast expansion in size since it can help researchers identify results from this collection that are appropriate for their needs. In addition to recommendation techniques like content-based suggestion and collaborative filtering, which do not let the user define their own needs, the focus is now on personalized recommendation platforms that can assess individual preferences by analyzing inputs. However, top-notch recommendation systems that meet each user's unique demands rely heavily on dataset assumptions. Our goal in this study is to demonstrate an on-demand recommendation system that can acknowledge how quickly research papers are expanding. The EIHI (Efficient Incremental High Utility items set mining algorithm), which is designed to cope with changing data, has recently been introduced in the compositions and is utilized for accomplishing the necessary. Empirical results demonstrate that the suggested approach effectively addresses the incremental nature of this analytic article while also satisfying the researcher's customized needs [4].

c) Knowledge-based Recommender System (KBRS)

Every time we look for a thing that we don't often buy, like a vehicle, bike, apartment, etc. The absence of adequate information regarding these search items makes recommendations challenging. Few other details are required for that, such as the site visited by the user behaviour. This system makes recommendations on the basis of a deeper understanding model that considers the relationship between the user's actions and the objects at hand. A fundamental component of KBRSs is the case-based reasoning technique, which divides a user's request into many characteristics based on numerous approaches and offers recommendations that are comparable to one's expected inclinations. Constraint-based recommendation system is another kind of KBRS that displays objects depending on one's preferences and adapts to the individual's proclivities. In a situation where a search result is not found, a group of related items is suggested instead. A diverse information base about consumers and items may be established with the use of e-commerce technologies. It makes use of ontologies, a formal knowledge representation of the process by which individual domain expertise and different aspects [5]. It is possible to gauge how comparable two items are by looking at each one's category. To demonstrate the correct match for the recommendation, data from a user overview and item description are used. Using an exegetic-based recommendation system eliminates a number of typical RS difficulties. Additional information on the exegetic-based recommendation system is available in the specified item as shown in Fig. 1. For a social site, for instance, the authors proposed and investigated the use of an exegetic-based recommendation system. We know that knowledge modelling and operating procedures in knowledge-based resource repositories (KBRSs) are somewhat expensive, but they are also capable of delivering all the required information that cannot be obtained in traditional ways [6].

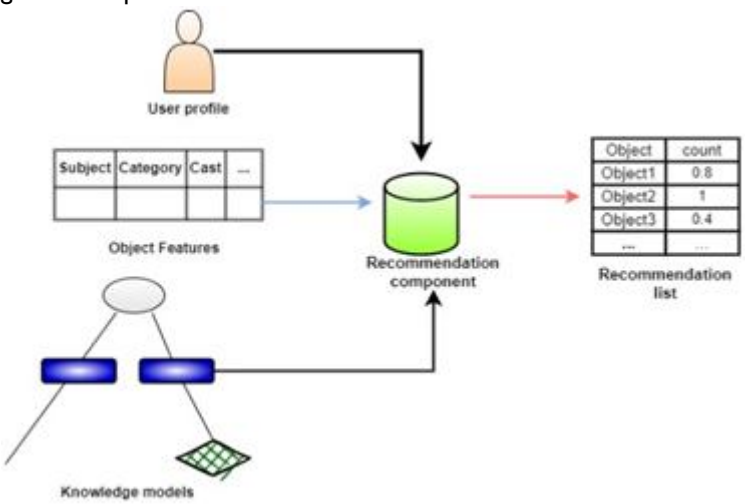

Fig.1. Knowledge-based recommender System

II. LITERATURE SURVEY

In machine learning, the K-means method is one of the most straightforward and effective clustering techniques. Although it is a commonly used technology among the current methods, the reliability of the outcomes may be diminished owing to chaotic information and outliers.

For this reason, we are using Hui Xu upgraded K-means clustering approach, where noisy information is removed using the clique matrix. The motivation behind K-implies calculation is to track down similitude between information by iteratively limiting the action between bunch focus and noticed information [7]. K-implies is a solo calculation that is utilized in bunching datasets. The system of recommendation collects data and generates an algorithm that takes into account the various demands of the consumers. Three approaches may be used to broadly categorize the system of recommendation. Hybrid, content-based, and collaborative approach to recommendation. Traditional recommendation systems do not take into account the surrounding surroundings, such as bus stops and metro stations, and instead suggest accommodations based only on price and rating [8]. The recommendation system provides the customer with appropriate lodging recommendations based on their choices for the nearby area. In order to provide useful service suggestions, the location of the user is essential data that may be connected to their current user profile. Furthermore, a significant number of GPS trajectories indicating user movement records are produced by the ubiquitous accessibility of GPS-enabled devices. It is possible to find intriguing usage trends using these GPS trajectories. We looked at the usefulness and applications of data obtained from user GPS trajectories in recommender systems. A system of recommendations makes suggestions for goods or services based on the user's past purchases or preferences [9].However, with current developments, recommendation systems make service recommendations based on geo location information. Additionally, it provides the user with more suitable options from the extensive database [10].

III. COLLABORATIVE FILTERING SYSTEM

To locate users with comparable features, the user can be classed in the collaborative filtering system according to the characteristics of their similar rating preference [11]. Predicting user preferences through the identification of several users' interests and information is known as collaborative filtering. This is accomplished by employing methods that need cooperation between several agents, data sources, etc. to filter data for information. The method then suggests products that these comparable users have chosen [12].

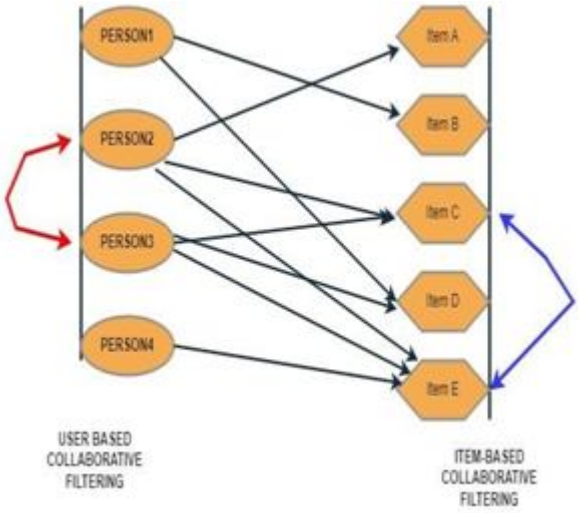

Fig.2. Collaborative filtering system

- **Advantages**
  The primary benefit of utilizing cooperative separating models is the effortlessness of execution and the elevated degree of inclusion they give. It likewise enjoys the benefit that it catches unobtrusive elements (extremely valid for dormant component models) and doesn't need a reasonable comprehension of the thing's items.
- **Disadvantages**
  Due to user/item interaction, this model's primary drawback is that it is not user-friendly when it comes to suggesting new products.This is known as the "cold start issue."

IV. MEMORY-BASED APPROACH

Memory-based Based Approach consistently called area cooperative separating. Significant, rankings of shopper object blends are normal on the reason in their areas. This might be also cut up into buyer basically based on absolutely cooperative separating and object principally based on absolutely cooperative separating. Client fundamentally based absolutely essentially way those similar clients will yield durable and tantamount suggestions. Things basically based on absolutely cooperative separating suggest protests fundamentally based absolutely on the likeness among objects determined the use of customer rankings of these items [13].The expectation method in memory-basically based absolutely CF comprises three stages. They are similitude appraisal, time of closest areas, and what's more, evaluating forecast. For evaluation of the general presentation, the CF gadget considers the suggested outright mistake (MAE), accuracy, and remembers. The CF generally speaking presentation shifts steadily with the handling approach of each and every step.

Recommendations are generated by the content based system according to the profiles and interests of the users. They attempt to match users with previously favoured goods. The characteristics of the components that the user rates highly are often used to determine the degree of similarity between the elements. The similarity between two articles is computed using the cosine similarity concept. Two vectors are created from the transformed dataset [14]. The ratio of the dot product to the product of the magnitudes of the two vectors, P and Q, is known as similarity.
If two vectors are alike, then their similarity will be one or very close to one. If two different vectors are exists, it will be either zero or very close to zero. Cosine similarity is followed as:

$$\text{Similarity (P,Q)} = \frac{P.Q}{||P||||Q||} = \frac{\sum_{i=1}^{n} P_i \times Q_i}{\sqrt{\sum_{i=1}^{n} P_i^2} \times \sqrt{\sum_{i=1}^{n} Q_i^2}}$$

Steps to find similarity between two articles:

- Obtain the article's title and index.
- Find the list of all the articles and the cosine similarity for the specific film.
- Transform the list into a tuple data structure, where the position is indicated by the first element and the similarity score is shown by the second.
- Sort the tuple list according to the similarity score.
- Ignore the first entry in the list of the top ten.
- The film itself is the one at the top with the greatest degree of similarity to the specific film.
- Thus, the most comparable film comes in second.

Advantages

The best time to use content-based models for item recommendations is when insufficient rating data is available. In fact, people may have rated other things with comparable features. Thus, even in the absence of a large amount of data, a model has to be able to use ratings in conjunction with item features to provide suggestions.

Disadvantages

The suggestions made are "obvious" given the products and material the user has viewed. This is a drawback as the user will never receive a recommendation for a certain kind of item if they have never dealt with it. For instance, using this method won't ever suggest mystery novels to you if you've never read any. This is due to the model's user-specificity and lack of utilization of knowledge from comparable users. As a result, there is less variation among the suggestions, which is bad for many firms.

V. HYBRID APPROACH

The goal of hybrid recommendation systems is to produce reliable conclusions by utilizing a variety of accessible data sources. There are two main concepts for hybrid recommendation systems: sequential and parallel [15]. Various systems of recommendations get input from the parallel design, and those suggestions are integrated to produce a single output. The specifications for an input are supplied by the sequencing layout [16]. Result is sequentially forwarded to the subsequent recommender from a single recommendation engine. Check out the following figure for a visual depiction of the two designs.

Types of Hybrid

1. **Weighted Hybrid**: In the weighted recommendation system that perform well in interpreting dataset. The outcomes from each model are combined by the weighted recommendation system to create a static weight that remains constant across the testing and training sets.
2. **Mixed Hybrid:** The union technique to generate several candidate information. Consequently, the recommendation system will incorporate several candidate groups into the model and merge the forecasts to generate the consequent advice.

Formulas used to recommend models:

$$\text{Final Score } (item) = \Sigma_i^n = 1(\text{Weight}i \times \text{Score}i(item))$$

where:

- n is the number of recommendation algorithms.
- Weight i is the weight assigned to the i-th algorithm.
- Score (item)is the score assigned to the item by the i-th algorithm.

Hybrid Recommendation System Algorithm:

1. **Data preparation:**
   - Gather user-item interaction data (user ratings, purchase history, etc.).
   - Prepare item features (content-based information) and user features if applicable.
2. **Train Individual recommendation algorithms:**
   - Train different recommendation algorithms, such as collaborative filtering, content-based filtering, and possibly others, on the prepared data.
   - Save the trained models for later use.

3. **Generate recommendations:**
   - For each user, use the trained models to generate recommendations independently.
   - Collaborative Filtering: Predict user-item interactions based on user behavior and preferences.
   - Content-Based Filtering: Recommend items similar to those the user has interacted with based on item features.
   - Other Recommendation Algorithms: Apply other algorithms based on your system's design.
4. **Recommendation Refinement:**
   - Apply post-processing steps if needed, such as removing already interacted items or ensuring diversity in recommendations [16].
5. **Return Final Recommendations:**
   - Provide the final list of recommendations to the user.

Advantages

Hybrid systems integrate many models to counteract the drawbacks of each individual model. In general, this minimizes the drawbacks of utilizing several models and helps provide suggestions that are stronger. Users receive suggestions that are more comprehensive and tailored as a result.

Drawbacks

For these models to remain current, a sizable database of ratings and other attributes is necessary, and they frequently have considerable computational difficulties. In the absence of current measurements (such as user interaction, ratings, etc.), it is challenging to retrain and offer fresh suggestions with updated items and evaluations from a range of users [16].

VI. PROPOSED METHODOLOGY

This initiative will gather the applicant's city details and use its data set, which includes all location related information. The gathered data will undergo additional cleansing to eliminate any unusual and erratic data. The K-means method will be used to cluster this processed data according to the user's location. We will be able to access any hostel and paid guest room in the area by using the Foursquare API [17]. Content-related filtering will be used to filter the generated dataset according to the client's preferences. Finally, maps will be used to show all of the generated data. The processes we use in our initiative to suggest hostels or private rooms to customers are as follows:

a) **Data collection:** This step will involve gathering the user-provided city's data. The city's data collection will be accessed in order to do this. Here, "data" refers to the specifics of each city location. The entirety of the city's info is contained in the data collection.

b) **Data Filtration:** In this phase, we will utilize the first stage's acquired data and clean it by removing any data that is noisy and anomalies that are not relevant to our goal. Pandas will be used for this. The removal of any unnecessary data makes the analysis more city-centric and will be ready for examination.

c) **Data Visualisation:** In this phase, the collected data will be shown on the graph as boxplots. Using Seaborn, Matplotlib, or Pandas, all of the cleansed data from the city will be shown.

d) **Fetching Geographical Information:** We will utilize the geopy API to retrieve the user's entered location, which gives us access to updated data for the designated spot. A popular Python package that permits developers to obtain user-entered geo location data via the Geopy API. It obtains extra information about the place and its user previous work, along with information of different locations, by accessing the user's location, separated by commas in both the north and south directions.

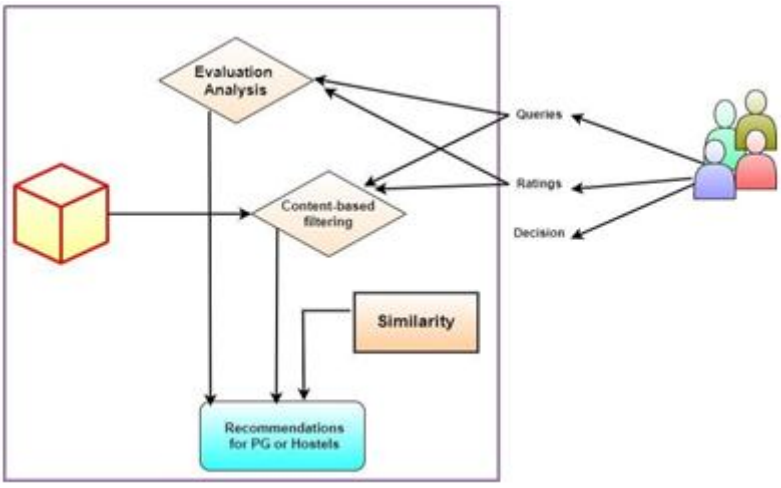

Fig.3. Recommendation based on user preferences

e) **Collection of Data:** The collected geo location data will be clustered in this stage, and a set of all the hostels and paying guest rooms will be returned, along with the hostels and paying guest data that are closest to the specified location. To accomplish this, we will utilize the K Means calculation, which is a bunching calculation, in view of closeness and we can utilize ScikitLearn.

f) ScikitLearn is an AI device, in this**;** there are different expectation calculations like Characterization, Bunching, and Relapse. In which we'll utilize Bunching, to group every one of the comparable information objects, in our undertaking we will bunch every one of the Lodgings and Paying Visitor Rooms information whose geo location is like the area given by the client [18].

g) Data visualization on maps: At this point, we will present the data that we gathered during the classification system on the guide. To do this, we will make use of the widely used Google Guides. We can use the Folium or Seaborn to present the findings in the guide. Folium and Seaborn are both used to visualize the data. Folium focuses more on the guide's perception, which is where data is initially presented. Using the longitude and scope of every point in the set along with its name, Folium will take the bunched set that was the output of K-Means and present its outcome on the guide.

h) Suggestion as per User Interest: In this case, we will utilize Content-dependent Filtering, which is a Natural Language Processing application in which the recommendation is dependent on what the user has supplied as input. This will be achieved by giving the customer a sifting option that they may select according on their preferences, just like when making selections about food, cost, and convenience [19]. The client's inclinations are listed in rundown form, and the Substance Based Separating will consider the convenience based on the client's choices. There have been numerous uses of Substance sifting like in Netflix, and film suggestions as shown in fig.3.

VII. CONCLUSION &FUTURE SCOPE

This research has concluded that anyone may publish the room for rental purposes. Selecting from a wide range of possibilities and based on the vast quantity of virtual data has been and always will be a challenging task. This sort of outcome has been attained by excellent study in previous years, and a variety of recommendation procedures and approaches have been demonstrated. The paper provides a summary of the various recommendation strategies knowledge-based, hybrid, demographic, content-based, collaborative, and context aware that are used in the recommendation system. Small content analysis, over-specialization, sparsity, scalability, analogy, acronyms, and black box problems are among the several problems encountered throughout the creation and implementation of the recommendation system. We firmly think that soon-to-be recommendation systems will not only benefit business, but also significantly affect our daily lives. These technologies will genuinely become ubiquitous and turn into indispensable instruments in every aspect of our life. The majority of recommendation systems developed to date have been targeted at vendors, buyers, and service providers; in other words, they are intended to draw in the poor. But recommendation systems in the future won't be limited to apps for buying and selling goods—instead, they will actually take action. Someone who knows us more intimately than we know ourselves should be an ideal RS, much like a close companion. Even in the event that we don't demonstrate precision, they need to be able to discern our needs and respond appropriately. The Internet of Things (IoT), artificial intelligence (AI), affective computing, cognitive computing, and other technologies will all have a significant impact in the future. By comprehending both users and items, it will be able to map users and items more effectively and be used in a broader range of applications. The line separating suggestions and searches will become less distinct in the future recommendation system. A key component of the search engine of the future that may provide tailored search will be the recommendation system. Customers find it difficult to understand the reasoning behind the advice that they purchase Herlocker. It will be more open for the RSS. For this reason, it matters if we are combining searching and recommending. People will search more carefully if they are aware of the pattern of recommendations. The recommendation system will be more user-friendly and will continuously enhance its content by incorporating input from a variety of sources. By supporting various reviews, it will become RSs to provide novel recommendation techniques ductile. Recurrent neural networks (RNNs) and will enable them to be precise in context; reinforcement learning will be extended by future time and emotion.

REFERENCES

1. Hu, Y., Yang, Y. and Huang, B. (2015) 'A comprehensive survey of recommendation system based on taxi GPS trajectory', in International Conference on Service Science.
2. Ignatov,D.I.,Kaminskaya A.,Konstantinova, N. and Konstantinov, A.Jun Wang, Arjen P. de Vries, Marcel J.T. Reinders," Unifying User based and Item based Collaborative Filtering Approaches by Similarity Fusion", 2006 ACM.
3. Panagiotis Symeonidis*, Alexandros Nanopoulos, Apostolos N. Papadopoulos,Yannis Manolopoulos," Collaborative recommender systems: Combining effectiveness and efficiency", 2007 Elsevier Ltd.
4. Kazuyoshi Yoshii, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, Hiroshi G. Okuno," An Efficient Hybrid Music Recommender System Using an Incrementally Trainable Probabilistic Generative Model",IEEE 2008.
5. Akhmed Umyarov, Alexander Tuzhilin, "Improving Collaborative Filtering Recommendations Using External Data",2008 IEEE.
6. Zunping Cheng, Neil Hurley," Effective Diverse and Obfuscated Attacks on Model-based Recommender Systems" 2009ACM.
7. Hendrik Drachsler, Hans G.K. Hummeland Rob Koper, "Personal recommender systems for learners in lifelong learning networks: the requirements, techniques and model".

8. Mohammad Yahya H. Al-Shamri, Nagi H. Al-Ashwal," Fuzzy-Weighted Similarity Measures for Memory-Based Collaborative Recommender Systems", Journal of Intelligent Learning Systems and Applications, 2014.
9. Recommender Systems: an Introduction, Cambridge University Press, New York.
10. Jäschke, R., Marinho, L.B., Hotho,A., Schmidt-Thieme, L. and Stumme, G. (2007) 'Tag recommendations in folksonomies',in 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland. Jinni (2017).
11. Chen, X. et al. (2018) Datamed–an open source discovery index for finding biomedical datasets. Journal of the American Medical Informatics Association, 25, 300–308.
12. Prasetyo, S. E., Utomo, A. B., & Hudallah, N. (2018). Implementation of Google Maps API 3 with Haversine Algorithm in the Development of Geographic Information System Boarding House Finder. In Proceedings of the 7th Engineering International Conference on Education, Concept and Application on Green Technology, Eic (pp. 227-233).
13. Rathore, K., Syed, A., & Patel, R. (2021). RENTAL HOUSE MANAGEMENT SYSTEM.
14. Hasnat, Ifrat, Nishat Tasnim, and Debabrata Chakraborty. "Bachelor House Rental Management System." (2019).
15. Gommans, H. P., Njiru, G. M., & Owange, A. N. (2014). Rental House Management System. International Journal of Scientific and Research Publications, 4(11), 1-24.
16. R. Saga, Yoshihiro Hayashi, and H. Tsuji. Hotel recommender system based on user's preference transition. In 2008 IEEE International Conference on Systems, Man and Cybernetics, pages 2437–2442, Oct 2008.
17. Priya Naik, Palak V. Desai, Supriya Pati. Location Based Place Recommendation using Social Network. IEEE 2019.
18. Koji Takuma, Junya Yamamoto, Sayaka Kamei, Satoshi Fujita. A hotel recommendation based on reviews: What do you attach important to IEEE 2016.
19. Recommendation System for Property Search Using Content Based Filtering Method. IEEE April 2018.

IRJCS: ©2014-24, AM Publications, India - All Rights Reserved