JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Hotel Recommendation System Using Machine Learning

Meduri V N S S R K Sai Somayajulu [1], E. Mahendra [2], B. Seshagiri [3],

S. Raghu Kumar [4], K. Bhanuji Rao [5]

[1] Assistant professor, Dept of CSE, Krishna University College Of Engineering And Technology, Machlipatnam, Krishna(Dt), AP, India.

[2] Final year student, Dept of CSE, Krishna University College Of Engineering And Technology, Machlipatnam, Krishna(Dt), AP, India.

[3] Final year student, Dept of CSE, Krishna University College Of Engineering And Technology, Machlipatnam, Krishna(Dt), AP, India.

[4] Final year student, Dept of CSE, Krishna University College Of Engineering And Technology, Machlipatnam, Krishna(Dt), AP, India.

[5] Final year student, Dept of CSE, Krishna University College Of Engineering And Technology, Machlipatnam, Krishna(Dt), AP, India.

Abstract: ***In today's digital age, recommendation systems play a crucial role in assisting consumers across various industries, most notably in entertainment and tourism. The proliferation of these systems has revolutionized the way travelers explore and select accommodations, particularly in the global hotel industry. By leveraging advanced algorithms and customer-generated data, recommendation systems help users make informed decisions by offering tailored suggestions that align with their preferences. This paper delves into the significance of recommendation systems in hotel selection, focusing on the challenges and advancements in online tourism resource recommendations.***

Keywords: SVM, Naïve Bayes, KNN.

I. INTRODUCTION

When it comes to creating recommendations for products, recommendation systems are an extremely important factor. They are used to filter information coming from a variety of networks and to forecast the output depending on the preferences of the user. These systems have gained a significant amount of popularity, and one business that might benefit from using recommender systems is the tourism industry. The use of recommendation systems has been shown to increase both the level of happiness and overall experience that customers have in a variety of businesses, including the travel industry [1]. They are generating enormous amounts of income using this method, which is why the vast majority of them are turning to recommendation systems [3]. A subfield of machine learning is called as recommendation engines, and its primary function is to usually rank individuals or items. A recommender system, in its broadest sense, is a system that predicts the ratings that a user would give to a certain item based on their previous interactions with that item. After that, the user will be presented with a rating of these different projections. In order to provide users with product recommendations that are likely to pique their interest, recommender systems make an effort to anticipate the preferences of users. In order to get the The purpose of a hotel recommendation system is to separate the user's selection of a recommended hotel or resort from their other available alternatives. The user's favorite hotel or resort should stand out among the other potential options in a hotel recommendation system, since that is the system's primary mission. If you wish to travel for work, for example, the hotel suggestion system should show you the lodgings that other customers have rated as being the best for business travel. Related Work

A huge quantity has already been done and written about item recommendation systems. A common method for those systems is to use a person-item matrix combining capabilities about the customers and gadgets together with person feedback for the gadgets. However, these methods proved to be inapplicable for our challenge, as the anonymized nature of the target variables made it tough to gain relevant features for them. extra applicable to our project, previous paintings has been performed on lodge recommendation systems with the aid of GAO Huming and LI Weili, who showed top consequences the use of a mixture of clustering and boosting algorithms. Even as their outcomes aren't similar to ours because of the large differences within the datasets used, it is great that each their paper and ours show promise in the usage of clustering and boosting for hotel recommendations.

II. LITERATURE REVIEW

J. Calero-Sanz, A. Orea-Giner, T. Villacé-Molinero, A. Muñoz-Mazón, and L. Fuentes-Moraleda, "Predicting A New Hotel Rating System by Analysing UGC Content from Tripadvisor: Machine Learning Application to Analyse Service Robots Influence," Procedia Comput. Sci., vol. 200, no. 2019, pp. 1078–1083, 2022.

Industry 4.0 tools permit computerized creation measures, and Artificial Intelligence (AI) approaches are pivotal in investigating the travel industry. Applying these devices to decipher User Generated Content (UGC) is fundamental to understand better client's necessities, opinions, and assumptions regarding tourism services. Through this research, an exploratory analysis of results is developed through Machine Learning Models to understand better the role played by robot and traveler typologies on the rating given to hotels considering TripAdvisor reviews of 74 hotels. The purpose of this exploratory research is to develop a methodology focused on analyzing online reviews related to service robots in hotels using Machine Learning techniques to train the data collected from TripAdvisor. Preliminary results show a link between the hotel rating given in TripAdvisor and the robot typology implemented in hotels.

N. Čumlievski, M. Brkić Bakarić, and M. Matetić, "A Smart Tourism Case Study: Classification of Accommodation Using Machine Learning Models Based on Accommodation Characteristics and Online Guest s Reviews," Electron., vol. 11, no. 6, 2022, doi: 10.3390/electronics11060913.

This paper deals with the analysis of data retrieved from a web page for booking accommodation. The main idea of the research is to analyze the relationship between accommodation factors and customer reviews in order to determine the factors that have the greatest influence on customer reviews. Machine learning methods are applied to the collected data and models that can predict the review category for those accommodations that are not evaluated by users are trained. The relationship between certain accommodation factors and classification accuracy of the models is examined in order to get detailed insight into the data used for model training, as well as to make the models more interpretable. The classification accuracy of each model is tested and the precision and recall of the models are examined and compared.

M. S. A. Forhad, M. S. Arefin, A. S. M. Kayes, K. Ahmed, M. J. M. Chowdhury, and I. Kumara, "An effective hotel recommendation system through processing heterogeneous data," Electron., vol. 10, no. 16, pp. 1–21, 2021, doi: 10.3390/electronics10161920.

Recommendation systems have recently gained a lot of popularity in various industries such as entertainment and tourism. They can act as filters of information by providing relevant suggestions to the users through processing heterogeneous data from different networks. Many travelers and tourists routinely rely on textual reviews, numerical ratings, and points of interest to select hotels in cities worldwide. To attract more customers, online hotel booking systems typically rank their hotels based on the recommendations from their customers. In this paper, we present a framework that can rank hotels by analyzing hotels' customer reviews and nearby amenities. In addition, a framework is presented that combines the scores generated from user reviews and surrounding facilities. We perform experiments using datasets from online hotel booking platforms such as TripAdvisor and Booking to evaluate the effectiveness and applicability of the proposed framework. We first store the keywords extracted from reviews and assign weights to each considered unigram and bigram keywords and, then, we give a numerical score to each considered

III. OBJECTIVE

This project is to enhance the effectiveness of online hotel recommendation systems in the tourism industry. Through a comprehensive literature analysis and the exploration of advanced

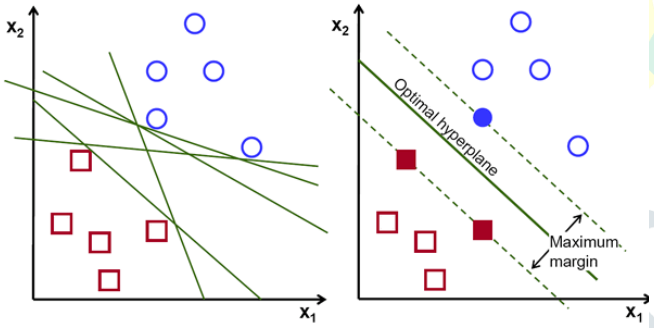
recommendation algorithms such as SVM, Naïve Bayes, and KNN, the project aims to address the challenges associated with recommending hotels based on crowdsourced data. Ultimately, the goal is to develop a model that empowers users to make informed hotel choices aligned with their preferences by leveraging customer feedback and improving the accuracy of hotel predictions.

IV. SCOPE

This project focuses on online hotel recommendation systems, specifically targeting the challenges associated with leveraging crowdsourced data for improving recommendation accuracy. The study encompasses the analysis and development of recommendation algorithms, including SVM, Naïve Bayes, and KNN, to better align hotel suggestions with user preferences. It considers the use of customer feedback as a crucial factor in the recommendation process.

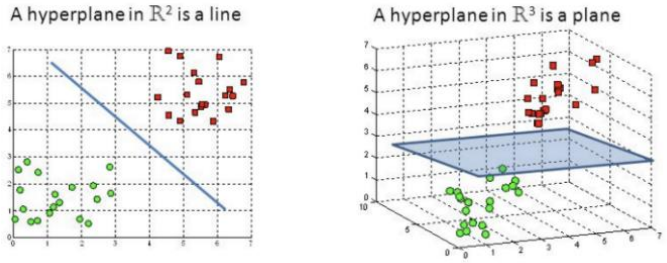
V. SUPPORT VECTOR MACHINES

support vector machine algorithm is to find a hyper plane in an N-dimensional space (N — the number of features) that distinctly classifies the data points.

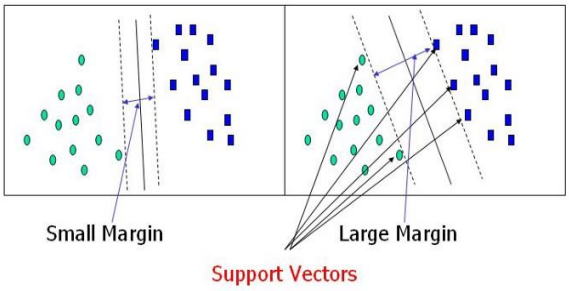


Possible hyper planes:
To separate the two classes of data points, there are many possible Hyper planes that could be chosen. Our objective is to find a plane that has the maximum margin, i.e. the maximum distance between data points of both classes. Maximizing the margin distance provides some reinforcement so that future data points can be classified with more confidence.

Hyper planes and Support Vectors

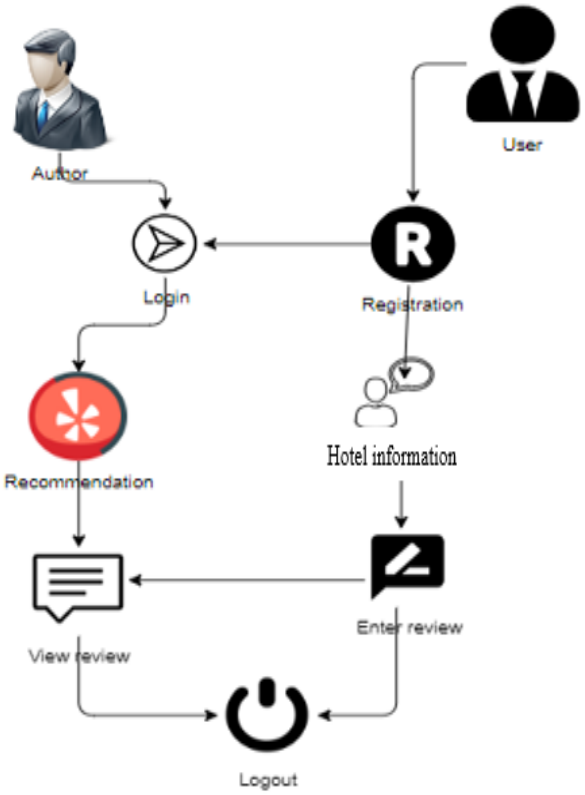


Hyper planes in 2D and 3D feature space
Hyper planes are decision boundaries that help classify the data points. Data points falling on either side of the hyper plane can be attributed to different classes. Also, the dimension of the hyper plane depends upon the number of features. If the number of input features is 2, then the hyper plane is just a line. If the number of input features is 3, then the hyper plane becomes a two-dimensional plane. It becomes difficult to imagine when the number of features exceeds 3.



Support Vectors
Support vectors are data points that are closer to the hyper plane and influence the position and orientation of the hyper plane. Using these support vectors, we maximize the margin of the classifier. Deleting the support vectors will change the position of the hyper plane. These are the points that help us build our SVM.

VI. ARCHITECTURE



VII. RESULTS AND DISCUSSION

Overall, The hotel recommendation system, powered by machine learning, effectively suggests personalized hotel options based on user preferences and booking history, significantly enhancing user satisfaction. Implementing a machine learning-based hotel recommendation system offers numerous advantages for the hospitality industry. By leveraging advanced algorithms, the system analyzes vast amounts of data, including user preferences and hotel attributes, to generate tailored recommendations. It employs collaborative, content-based, or hybrid filtering approaches to ensure accurate suggestions. Additional features such as location-based recommendations and sentiment analysis further enhance the system's performance.Moreover, the system continuously improves through machine learning, refining its recommendation algorithms based on user feedback. This adaptability leads to better personalized suggestions and increased user satisfaction.

VIII. CONCLUSION

Embarking on an innovative journey, our study pioneers the theoretical evaluation of hotel recommendation systems, employing the K-Nearest Neighbors (KNN) algorithm within content-based filtering. Breaking new ground, we intricately analyze the theoretical foundations of KNN's application, unraveling its nuances in the hotel domain. By seamlessly integrating content-based strategies, our approach surpasses conventional models, elevating recommendation precision. The ultimate novelty lies in predicting unique URLs for recommended hotels, introducing a tailored experience for users. This groundbreaking fusion of KNN and content-based filtering not only refines the theoretical understanding but also propels the practicality of hotel recommendations, marking a significant stride in personalized hospitality services.

IX. REFERENCES

[1] J. Calero-Sanz, A. Orea-Giner, T. Villacé-Molinero, A. Muñoz-Mazón, and L. Fuentes-Moraleda, “Predicting A New Hotel Rating System by Analysing UGC Content from Tripadvisor: Machine Learning Application to Analyse Service Robots Influence,” Procedia Comput. Sci., vol. 200, no. 2019, pp. 1078–1083, 2022, doi: 10.1016/j.procs.2022.01.307.

[2] N. Čumlievski, M. Brkić Bakarić, and M. Matetić, “A Smart Tourism Case Study: Classification of Accommodation Using Machine Learning Models Based on Accommodation Characteristics and Online Guest sReviews,” Electron., vol. 11, no. 6, 2022, doi: 10.3390/electronics11060913.

[3] J. Wu, C. Liu, Y. Wu, M. Cao, and Y. Liu, “A Novel Hotel Selection Decision Support Model Based on the Online Reviews from Opinion Leaders by Best Worst Method,” Int. J. Comput. Intell. Syst., vol. 15, no. 1, 2022, doi: 10.1007/s44196-022-00073-w.

[4] W. Zkr et al., “$Vshfw %Dvhg 6Hqwlphqw 2Ulhqwhg +Rwho 5Hfrpphqgdwlrq 0Rgho ([Sorlwlqj 8Vhu 3Uhihuhqfh /Hduqlqj,” no. 978, pp. 409–414, 2020.

[5] M. S. A. Forhad, M. S. Arefin, A. S. M. Kayes, K. Ahmed, M. J. M. Chowdhury, and I. Kumara, “An effective hotel recommendation system through processing heterogeneous data,” Electron., vol. 10, no. 16, pp. 1–21, 2021, doi: 10.3390/electronics10161920.

[6] M. M. Díaz and C. M. Duque, "Open innovation through customer satisfaction: A logit model to explain customer recommendations in the hotel sector," J. Open Innov. Technol. Mark. Complex., vol. 7, no. 3, 2021, doi: 10.3390/joitmc7030180.

[7] Y. J. Kim and H. S. Kim, "The Impact of Hotel Customer Experience on Customer Satisfaction through Online Reviews," Sustain., vol. 14, no. 2, pp. 1–13, 2022, doi: 10.3390/su14020848.

[8] M. Zhao, L. Li, and Z. Xu, "Study on hotel selection method based on integrating online ratings and reviews from multi-websites," Inf. Sci. (Ny)., vol. 572, pp. 460–481, 2021, doi: 10.1016/j.ins.2021.05.042.

[9] K. Al Fararni, F. Nafis, B. Aghoutane, A. Yahyaouy, J. Riffi, and A. Sabri, "Hybrid recommender system for tourism based on big data and AI: A conceptual framework," Big Data Min. Anal., vol. 4, no. 1, pp. 47–55, 2021, doi: 10.26599/BDMA.2020.9020015.

[10] M. S. Satu, K. Ahammed, and M. Z. Abedin, "Performance Analysis of Machine Learning Techniques to Predict Hotel booking Cancellations in Hospitality Industry," ICCIT 2020 - 23rd Int. Conf. Comput. Inf. Technol. Proc., pp. 19–21, 2020, doi: 10.1109/ICCIT51783.2020.9392648.

[11] M. S. A. Forhad, M. S. Arefin, A. S. M. Kayes, K. Ahmed, M. J. M. Chowdhury, and I. Kumara, "An effective hotel recommendation system through processing heterogeneous data," Electron., vol. 10, no. 16, pp. 1–21, 2021, doi: 10.3390/electronics10161920.

[12] B. Kaya, "A hotel recommendation system based on customer location: a link prediction approach," Multimed. Tools Appl., vol. 79, no. 3–4, pp. 1745–1758, 2020, doi: 10.1007/s11042-019-08270-0.

[13] S. Wei and H. S. Kim, "Online Customer Reviews and Satisfaction with an Upscale Hotel: A Case Study of Atlantis, The Palm in Dubai," Inf., vol. 13, no. 3, 2022, doi: 10.3390/info13030150.