HOTEL RECOMMENDATION SYSTEM USING MACHINE LEARNING WITH COMMISION BASED REVENUE MODEL

Sameer

October 28, 2023

Abstract

In this prototype report, a Hotel Recommendation System with a specific focus on its intricate technical model and user-centric features is presented. The system includes user registration, enabling personalized experiences, and showcases a user-friendly interface for viewing comprehensive hotel information. The heart of the prototype lies in the recommendation algorithm, which employs a hybrid approach, combining content-based and collaborative filtering techniques. This unique model processes a rich dataset that incorporates user preferences, hotel attributes, historical bookings, and reviews. It continuously adapts and learns from user interactions, ensuring that recommendations remain relevant.

Furthermore, the report delves into the performance evaluation, emphasizing the use of industry-standard metrics like Precision at K (P@K), Recall at K (R@K), and Mean Average Precision (MAP) to assess the system's recommendation quality. This prototype serves as a vital step in developing a user-centric platform concept that effectively combines data-driven recommendations with commission-based revenue generation.

1.0 Introduction

In today's fast-paced world of travel, one of the most common dilemmas travelers face is the overwhelming abundance of hotel options. It's like standing in front of an all-we-can-eat buffet with so many delicious choices that it's hard to decide where to start. The sheer number of hotels can lead to a problem known as "decision overload," where travelers often end up making subpar hotel choices due to sheer exhaustion from the selection process. This challenge is what we're determined to address with our prototype for a Hotel Recommendation System. Our goal is to simplify the hotel booking process and provide personalized recommendations so we can find the perfect place to stay without all the stress. In this report, we'll introduce our solution and discuss how it can make wer travel planning easier while also explaining our commission-based revenue model to keep this service running smoothly.

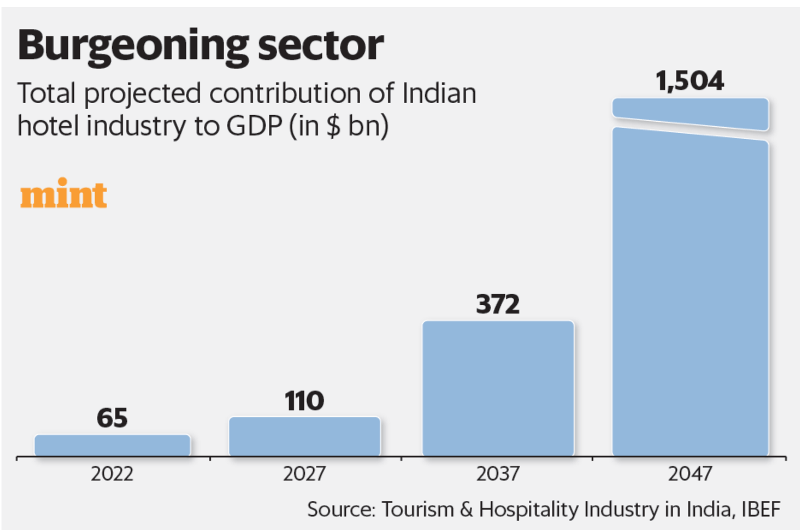
1.1 Problem Statement

This project seeks to build a machine learning-based hotel recommendation system, addressing the challenge of choice overload faced by travelers. The system will provide personalized recommendations, simplifying the hotel selection process, while a commission-based monetization model will ensure its sustainability and financial viability.

2.0 Business Need Assessment for Hotel Recommendation System

With a projected growth rate of 9.7% from 2023 to 2029, the Indian hotel market is set to expand substantially, reaching nearly USD 56.21 million. Furthermore, the hotels market is anticipated to generate USD 7.66 billion in revenue by 2023, with a notable compound annual growth rate (CAGR) of 8.28% from 2023 to 2027, indicative of the industry's robust prospects.

In parallel, the hospitality sector in India is poised for impressive growth, with an expected value of USD 23.5 billion in 2023. It is projected to maintain a healthy CAGR of 4.73%, eventually reaching USD 29.61 billion by 2028. These statistics demonstrate a burgeoning industry with significant economic potential, providing an ideal environment for the development and implementation of innovative solutions like a Hotel Recommendation System, poised to cater to the evolving needs and preferences of both domestic and international travelers.



3.0 External Search (Information Sources)

Selection of a suitable hotel location and reservation of accommodation have become a critical issue for the travelers. The online hotel search has been increased at a very fast pace and became very time-consuming due to the presence of huge amount of online information. Recommender systems (RSs) are getting importance due to their significance in making decisions and providing detailed information about the required product or a service. Hotel recommendation systems leverage vast amounts of data, including user preferences, hotel features, historical bookings, and user reviews.

Machine learning algorithms analyze this data to provide personalized recommendations to users. Machine learning models create user profiles based on their past interactions and preferences. These profiles are used to tailor hotel recommendations to individual users, ensuring a more personalized experience.

Content-based filtering considers the characteristics of both hotels and users. It matches user preferences with hotel attributes like location, price, amenities, and reviews to suggest the most suitable options.

Collaborative filtering identifies users with similar preferences and recommends hotels based on the choices of users with comparable tastes. This method enhances recommendations by considering the wisdom of the crowd.

Many recommendation systems use hybrid models that combine both content-based and collaborative filtering approaches. This combination aims to provide a more comprehensive and accurate recommendation.

3.1 Applicable Regulations

a.) Personal Data Protection Bill (PDPB): India is in the process of implementing the PDPB, which is expected to govern the collection, storage, and processing of personal data. It will require user consent, data protection measures, and data localization, among other provisions.

b.) Information Technology Act, 2000: This act addresses various aspects of electronic commerce and digital signatures. It includes provisions related to data protection, privacy, and cybersecurity.

c.) Payment Card Industry Data Security Standard (PCI DSS): If wer system handles payment information, it must comply with PCI DSS to ensure secure payment processing.

d.) E-commerce Regulations: Various e-commerce regulations may apply, especially if the platform enables hotel bookings and payments. Compliance with pricing transparency, consumer protection, and dispute resolution rules is essential.

e.) User Consent and Data Portability: Ensure that users are informed about data collection and processing practices. Provide options for users to control their data and request data portability.

3.2 Business opportunity

Travelers are increasingly seeking personalized travel services and recommendations. The market demand for such services is on the rise, making this an opportune time to enter the space.

By providing users with tailored recommendations that align with their preferences, the system can boost hotel bookings through the platform. This leads to increased revenue potential.

Offering a sophisticated recommendation system can give wer platform a competitive edge in the crowded travel and booking industry. It can differentiate wer platform from competitors.

Machine learning models can be scaled to handle increasing user demands and data volumes as the business grows.

The model can be extended to other travel-related services, such as restaurant recommendations, activity bookings, and transportation options.

4.0 Concept Generation

The concept of a hotel recommendation model using machine learning emerged from a recognition of the need for personalized, efficient, and data-driven solutions in the hospitality and travel industry. This concept was motivated by the following factual factors:

Market Demand: Extensive research and market analysis revealed a growing demand for personalized travel services. Travelers increasingly seek tailored recommendations for hotels that align with their preferences, making the concept of a recommendation system more relevant and market-driven.

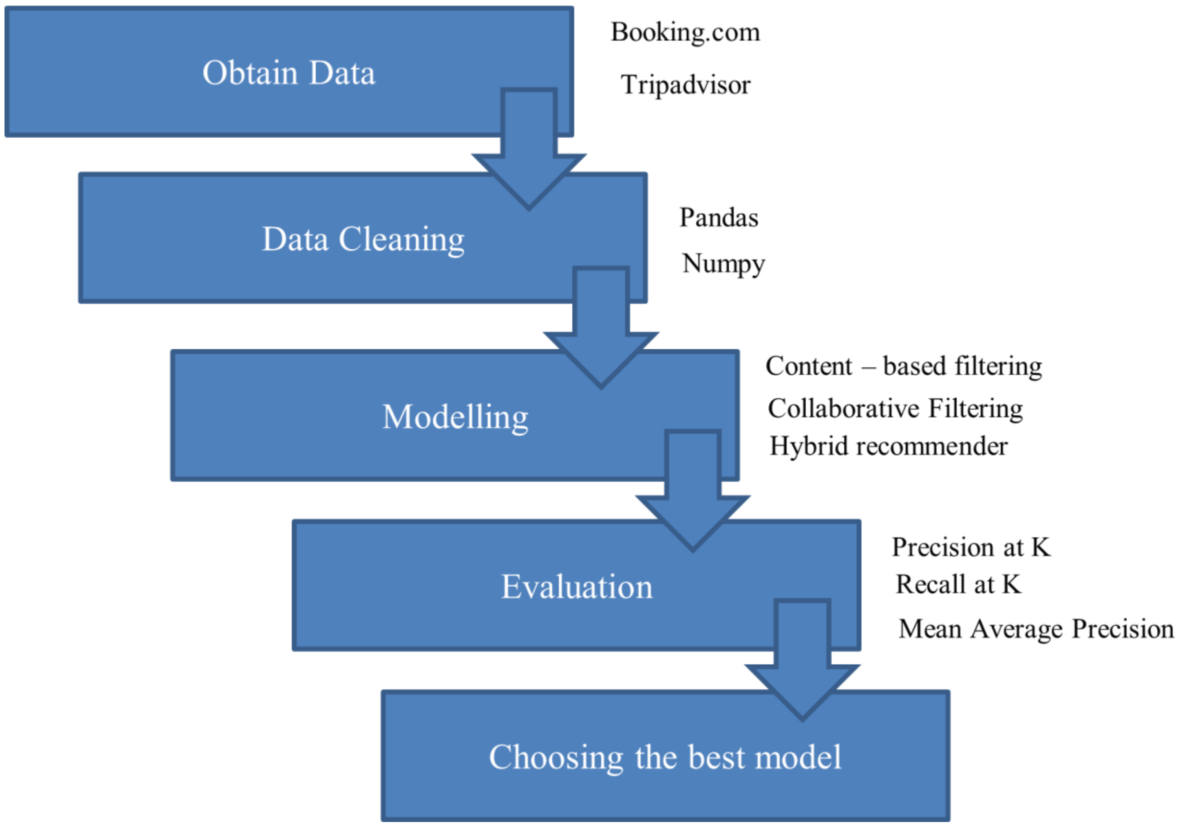
Advancements in Machine Learning: The rapid advancements in machine learning and artificial intelligence offered an opportunity to leverage these technologies for enhancing the travel and hospitality experience. Machine learning's ability to process vast datasets and discern intricate patterns was a key driver.

Data Availability: The availability of diverse and extensive data sources, such as user behavior, hotel attributes, and reviews, presented an opportunity to develop a model that could harness this data to generate valuable recommendations.

Parallel Trends: Observing the success of personalized, predictive models in other industries, such as healthcare, reinforced the notion that a similar approach could benefit the hotel and hospitality sector. This fact-based insight underscored the timeliness and relevance of the concept.

In essence, the concept of a hotel recommendation model using machine learning was driven by market demand, technological advancements, user-centric goals, data availability, and the alignment with trends in personalized and predictive services. These factual factors converged to inspire the development of a solution that enhances the traveler's hotel selection process.

4.1 Concept Development



Data Acquisition: Inititally we start by collecting data from multiple sources, including Booking.com and TripAdvisor. This data likely contains information about hotels, user interactions, reviews, and other relevant details. Combining data from various sources can enrich the recommendation system with a diverse set of information.

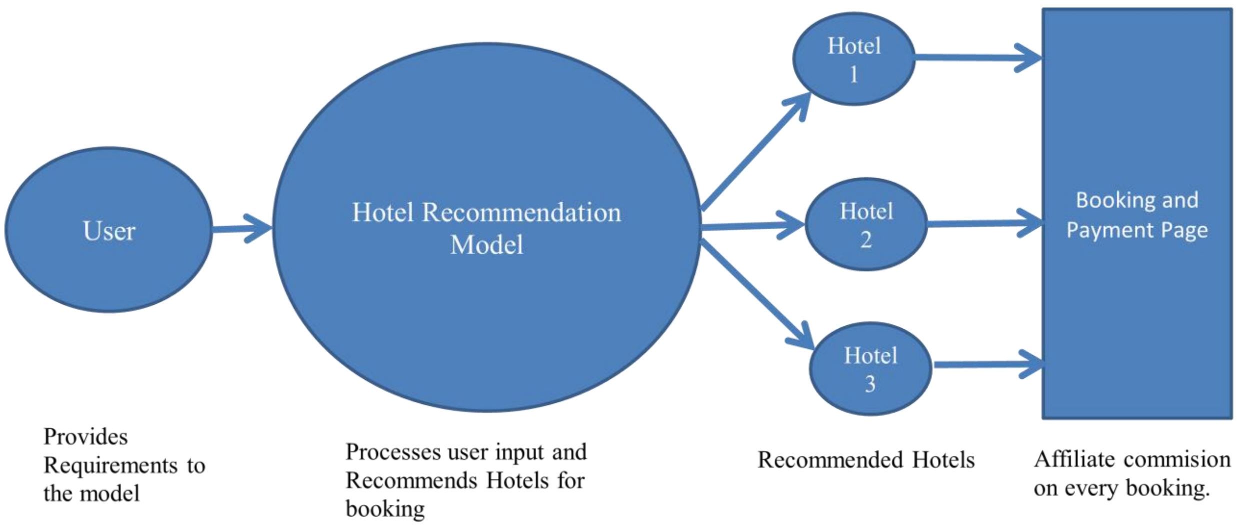
Data Cleaning and Preprocessing: Data preprocessing is crucial to ensure the quality and consistency of the data. In this step, we clean the data by handling missing values, removing duplicates, and addressing any inconsistencies. We also transform and structure the data to make it suitable for modeling. This might include text processing, numerical scaling, and creating user-item interaction matrices.

Modeling: We employ three recommendation techniques: content-based filtering, collaborative filtering, and a hybrid recommender. Content-based filtering leverages hotel characteristics to make recommendations, while collaborative filtering uses user interactions to identify similar users and items. The hybrid model combines the strengths of both approaches to provide more accurate and diverse recommendations. Modeling involves building and training machine learning models based on these techniques.

Evaluation: To assess the performance of wer recommendation models, we employ evaluation metrics such as Precision at K (P@K), Recall at K (R@K), and Mean Average Precision (MAP). These metrics help measure the accuracy, coverage, and quality of wer recommendations. By evaluating the models using these metrics, we gain insights into their strengths and weaknesses.

Model Selection: Based on the evaluation results, we choose the best-performing model. The model that excels in terms of P@K, R@K, and MAP may vary depending on wer specific goals and the characteristics of wer data. The chosen model should strike a balance between precision, recall, and overall recommendation quality.

4.2 Revenue Model:



User Interaction and Hotel Selection: Users visit the website and provide their hotel requirements, preferences, and constraints, such as location, price range, amenities, and travel dates. The recommendation system processes this input and generates a list of the most suitable hotels based on the user's criteria.

Hotel Booking: Users review the recommended hotels, read descriptions, view photos, and check reviews and ratings. Users select a hotel from the recommendations and proceed to the booking and reservation process.

Affiliate Commission: When a user successfully books a hotel through the website's booking platform, an affiliate commission is earned. The commission is a percentage of the total booking value and is typically paid by the hotel or a third-party booking service.

Revenue Generation: Revenue is generated based on the total affiliate commissions earned through successful hotel bookings made by users via the website.

5.0 Final Report Prototype

The product takes the following functions to perfect and provide a good result.

Back-end

Model Development: This must be done before releasing the service. A lot of manual supervised machine learning must be performed to optimize the automated tasks.

1. Performing EDA to realize the dependent and independent features.

2. Algorithm training and optimization must be done to minimize overfitting of the model.

Front End

1. Different user interface: The user must be given many options to choose form in terms of parameters. This can only be optimized after a lot of testing and analysis all the edge cases.

2. Interactive visualization: The data extracted from the trained models will return raw and inscrutable data. This must be present in an aesthetic and an “easy to read” style.

3. Feedback system: A valuable feedback system must be developed to understand the customer’s needs that have not been met. This will help us train the models constantly.

6.0 Code Implementation/Validation on Small Scale

6.1 Basic EDA

Data Cleaning and transformations

[13]:
```
del hotel_details['id']
del hotel_rooms['id']
del hotel_details['zipcode']
```

[14]:
```
hotel_details=hotel_details.dropna()
hotel_rooms=hotel_rooms.dropna()
```

[15]:
```
hotel_details.drop_duplicates(subset='hotelid',keep=False,inplace=True)
hotel=pd.merge(hotel_rooms,hotel_details,left_on='hotelcode',right_on='hotelid',how='inner')
```

[17]:
```
hotel.columns
```

[17]:
```
Index(['hotelcode', 'roomamenities', 'roomtype', 'ratedescription', 'hotelid',
       'hotelname', 'address', 'city', 'country', 'propertytype', 'starrating',
       'latitude', 'longitude', 'Source', 'url', 'curr'],
```

```
data_types = hotel.dtypes
data_types
```

[42]:
```
hotelcode          0
roomamenities      0
roomtype           0
ratedescription    0
hotelname          0
address            0
city               0
country            0
propertytype       0
starrating         0
latitude           0
longitude          0
dtype: int64
```

```
null_counts = hotel.isnull().sum()
null_counts
```

[37]:
```
hotelcode          0
roomamenities      0
roomtype           0
ratedescription    0
hotelname          0
address            0
city               0
country            0
propertytype       0
starrating         0
latitude           0
longitude          0
dtype: int64
```

6.2 Data after basic EDA

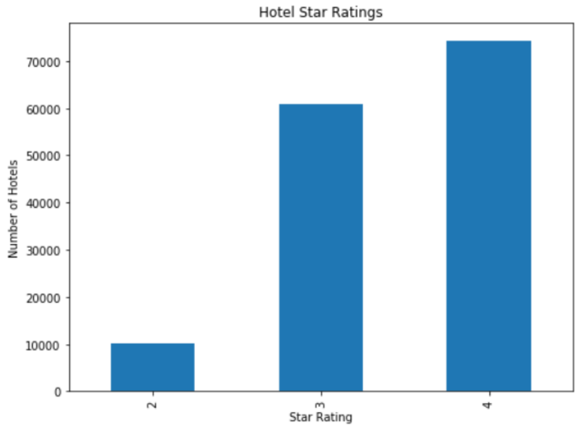
[33]:

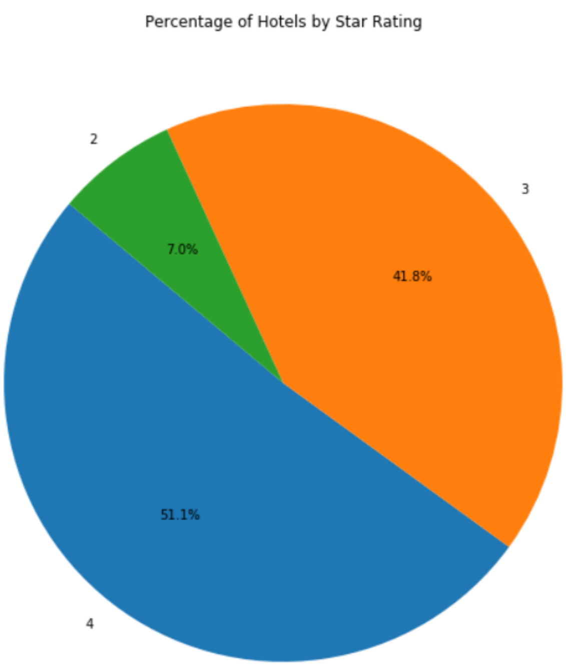
```
hotel.head()
```

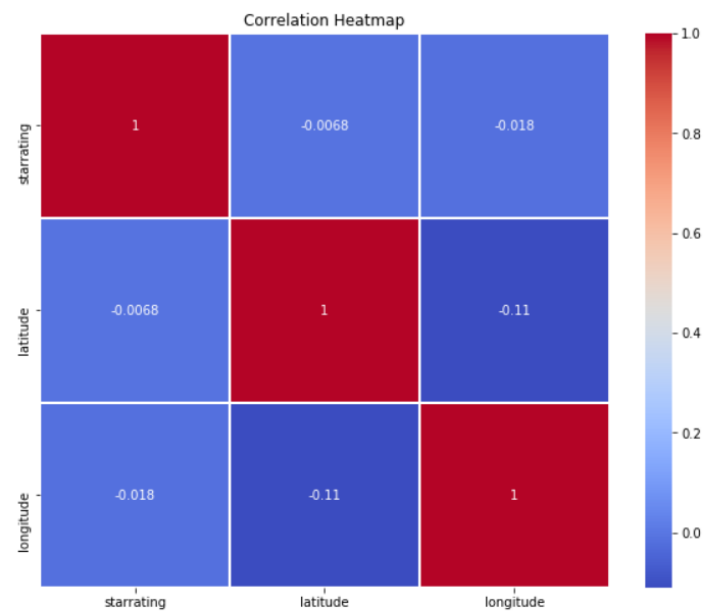
[33]:

	hotelcode	roomamenities	roomtype	ratedescription	hotelname	address	city	country	propertytype	starrating	latitude	longitude
0	634876	Air conditioning: ;Alarm clock: ;Carpeting: ;C...	Double Room	Room size: 15 m^2/161 ft^2, Shower, 1 king bed	The Old Cider House	25 Castle Street	Nether Stowey	United Kingdom	Hotels	4	51.150921	-3.15847
1	634876	Air conditioning: ;Alarm clock: ;Carpeting: ;C...	Double Room	Room size: 15 m^2/161 ft^2, Shower, 1 king bed	The Old Cider House	25 Castle Street	Nether Stowey	United Kingdom	Hotels	4	51.150921	-3.15847
2	634876	Air conditioning: ;Alarm clock: ;Carpeting: ;C...	Deluxe Double Room with Shower	Room size: 17 m^2/183 ft^2, Shower, 1 queen bed ...	The Old Cider House	25 Castle Street	Nether Stowey	United Kingdom	Hotels	4	51.150921	-3.15847
3	634876	Air conditioning: ;Alarm clock: ;Carpeting: ;C...	Superior Double Room	Room size: 17 m^2/183 ft^2, Shower, 1 double bed	The Old Cider House	25 Castle Street	Nether Stowey	United Kingdom	Hotels	4	51.150921	-3.15847
4	634876	Air conditioning: ;Alarm clock: ;Carpeting: ;C...	Standard Double or Twin Room	Room size: 13 m^2/140 ft^2, Shower, 1 queen bed ...	The Old Cider House	25 Castle Street	Nether Stowey	United Kingdom	Hotels	4	51.150921	-3.15847

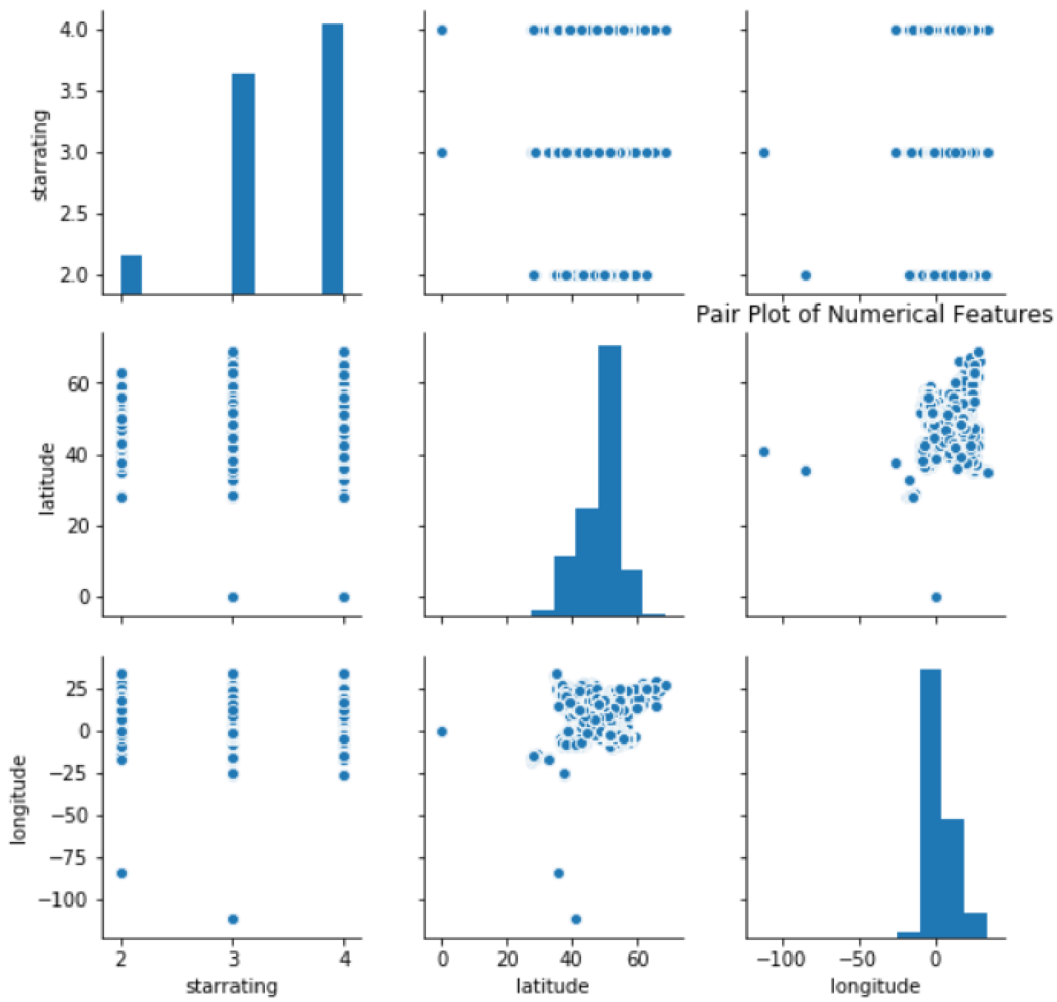
6.3 Data Visualization







Correlation Heatmap
1
-0.0068
-0.018
-0.0068
1
-0.11
-0.018
-0.11
1
starrating
latitude
longitude
starrating
latitude
longitude
1.0
0.8
0.6
0.4
0.2
0.0


Pair Plot of Numerical Features
starrating
latitude
longitude
4.0
3.5
3.0
2.5
2.0
60
40
20
0
25
0
−25
−50
−75
−100
2
3
4
0
20
40
60
−100
−50
0

6.4 ML Modelling

Recommender system based only on City and ratings about the hotel

```
[ ]:
def citybased(city):
    hotel['city']=hotel['city'].str.lower()
    citybase=hotel[hotel['city']==city.lower()]
    citybase=citybase.sort_values(by='starrating',ascending=False)
    citybase.drop_duplicates(subset='hotelcode',keep='first',inplace=True)
    if(citybase.empty==0):
        hname=citybase[['hotelname','starrating','address','roomamenities','ratedescription']]
        return hname.head()
    else:
        print('No Hotels Available')
```

6.5 Links

Github: Hotel Recommendation Using machine learning code

Kaggle Dataset: Hotel Recommendation Dataset

7.0 Conclusion

In the world of abundant hotel choices, our Hotel Recommendation System is the key to stress-free bookings, offering tailored suggestions. With our commission-based approach, we ensure both convenience and sustainability in travel planning.

The Indian hotel market is booming, with significant growth expected. This is a great opportunity for our Hotel Recommendation System to serve both local and international travelers with personalized solutions that match their preferences.

Choosing a hotel can be overwhelming with so many options online. Machine learning helps by giving personalized suggestions based on wer preferences. It makes hotel hunting easier and more enjoyable.

The idea for a hotel recommendation model using machine learning came from recognizing the increasing demand for personalized travel services, advances in technology, and the wealth of available data. These factors converged to create a user-centric solution for easier and more enjoyable hotel choices.

In creating a hotel recommendation system, we gather data from various sources, clean it, build models using content-based, collaborative, and hybrid filtering, and evaluate their performance using metrics like P@K and R@K. The best-performing model is selected based on its ability to provide accurate and diverse recommendations.

Users input their hotel preferences, select a hotel from recommendations, and complete bookings. The website earns revenue through affiliate commissions from successful bookings, ensuring sustainability and user convenience.

8.0 References

Ramzan, B., Bajwa, I.S., Jamil, N., Amin, R.U., Ramzan, S., Mirza, F. and Sarwar, N., 2019. An intelligent data analysis for recommendation systems using machine learning. Scientific Programming, 2019, pp.1-20.

Abbasi, F., Khadivar, A. and Yazdinejad, M., 2019. A grouping hotel recommender system based on deep learning and sentiment analysis. Journal of Information Technology Management, 11(2).

Arruza, M., Pericich, J. and Straka, M., 2016. The automated travel agent: hotel recommendations using machine learning.

Portugal, I., Alencar, P. and Cowan, D., 2018. The use of machine learning algorithms in recommender systems: A systematic review. Expert Systems with Applications, 97, pp.205-227.

Pazzani, M.J. and Billsus, D., 2007. Content-based recommendation systems. In The adaptive web: methods and strategies of web personalization (pp. 325-341). Berlin, Heidelberg: Springer Berlin Heidelberg.

Marlin, B., 2004. Collaborative filtering: A machine learning perspective (pp. 2239-2239). Toronto: University of Toronto.

Thakkar, P., Varma, K., Ukani, V., Mankad, S. and Tanwar, S., 2019. Combining user-based and item-based collaborative filtering using machine learning. In Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2018, Volume 2 (pp. 173-180). Springer Singapore.