Travel with nature and Analysis of Comments

Kirti Jain , Atul Singh , Bhasker Upadhyay , Harsh Dwivedi , Harsh Vardhan Singh

Department of Computer Science and Engineering, Inderprastha Engineering College, Ghaziabad, India

Abstract: As quoted by Gandhi Ji "Th future of India lies in its village", So many virgin beaches, undiscovered villages, and mountains gracing our land, there is still a lot to unravel before us and the world. The research will be focusing on those unexplored and undiscovered regions and promote tourism in those areas which will help them grow economically. We will also be analysing the comments posted by our visitors as positive, negative, or neutral to help our new visitors to get an idea about that place.

1. INTRODUCTION

Travellers in India are more focused on urban and luxurious places rather than rural areas, besides these, there are many unexplored regions in India or regions about which very few knows. Since very few know about these places they are yet underdeveloped and people are not getting the opportunities that they deserve.

Tourism in India is more about luxurious destinations and less about rural India. We will be highlighting the natural wealth of India because that's what actual India looks like. It will help the overall development of that area which will help to develop our economy. Rural Tourism can revitalize local art and crafts and prevent viable traditional occupations from being displaced. It will help redevelop rural areas and rejuvenate rural life. The interaction with the visitors will expand their knowledge and horizons.

Couch surfing which is a very new concept will be introduced so that interested travelers can have a different experience. Machine learning /Artificial Intelligence will be used to analyze people's comments which will give a better picture to the new visitors. It will help us to explore the unseen and unexplored regions of India. With so many virgin beaches, undiscovered villages, and mountains gracing our land, there is still a lot to unravel before us and the world. It will lead to enhanced tourism in many remote and rural areas which will lead to better business opportunities there and thus overall development of such places.

Sentiment analysis will help in retrieving information about a consumer's perception of a product, service or brand. We will be analyzing comments just to help travelers know more about the place just by rating them with positive,negative or neutral tag.Analyzing the opinion of the visitor and then differentiating between positive and negative aspects of places may help travelers to get an idea about how and what the place actually is.

2. RELATED WORK

The Popularity of the youtube video is analyzed after relating it to the sentiments of users comments[1] where supervised machine learning model is used.The proposed idea of cosine similarity measures the performance well for two and three class problem than Naive Bayes and SVM but leaves behind by a huge margin for five class classification[2]. Whereas, by combining the sentiment values of the tweet done by the user is used[3] to get to know about how user perceive and feel about the brand and services of the Tourism Industry. Sentiment analysis techniques and Python Library(Tweepy and TestBlob) is used. Naive Bayes, SVM and Lexicon Based approach has been used to classify the comments of the web services provided by the Tourism Industry[4]. Most of the comments are being classified as positive comments and didn't work well with neutral comments.

The Research Paper[5] studies the affect of using Classifiers (Naive Bayes, SVM and Genetical Algorithm) in hybrid and individual form. While dealing individually the accuracy is lesser as compared to the hybrid form. The polarity of the tweets are being checked as (positive, negative and neutral)[6] done by the user to get to know the user review of the Airline Services.

Support vector machine model is used[7] to classify the Hotel reviews as positive and negative to get to know the feedback better.The Research Paper[8] focused on Tourism in Unexplored region of Zimbabwe presented the idea of Cultural Tourism and development on economy and rural community. Similarly the research done in paper[9] has presented the

fact of increasing the tourist flow in West Bengal rural areas paving way for women empowerment and economic development.

Lexicon Based approach is used to measure the level of teachers from student comments[10]. Here, we categorize the student's feedback comments to strongly negative, moderate negative, weakly negative, strongly positive, moderately positive, weak positive, neutral category using two lexicons.

3 METHODOLOGY

To promote tourism in rural and natural India, our work will be mainly focused on those unexplored regions including Couchsurfing. For handling the backend we will be using the Django framework which is an open-source python framework for web development, also we will be using PostgreSQL for handling data.

Travelers will be adding a review about places and we will be analyzing and showing top positive and negative comments about those places which will help travelers to know more about the place. For doing so, we will be using NLTK (Natural Language Toolkit) Library which is a suite that contains libraries and programs for statistical language processing. It is one of the most powerful NLP libraries, which contains packages to make machines understand human language and reply to it with an appropriate response.

3.1 FRONTEND

For the frontend part, we used HTML, CSS, JAVASCRIPT, and SWIPER (fig-1).

HTML and CSS are used to design the website and beautify it respectively web content like applications and browsers, With the help of javascript we made our website interactive and dynamic, and with the help of Swiper responsive slider is created on the website.
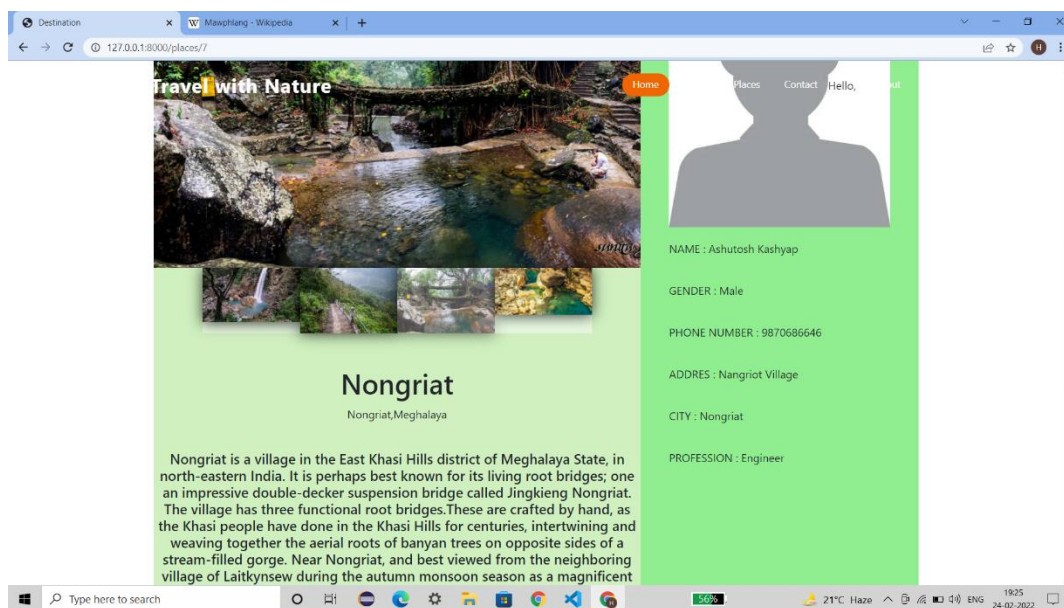

Fig(1)

3.2 BACKEND

The Backend was done Using the Django framework (fig-2).Django is an open-source python web framework.The user needs to login, in order to view or add comments in the destination section.The data of the user and comments are stored in the database (fig-3), We have used PostgreSQL for storing and retrieving data, the tool used to manage PostgreSQL is pgAdmin4.For logging in the user must enter their Name email and password. The password then is stored in the database in encrypted form, SHA-256 Hashing algorithm is used for the encryption of the password.For adding, deleting, and modifying places, the Admin Section has been created which is a User Interface to deal with backend information.
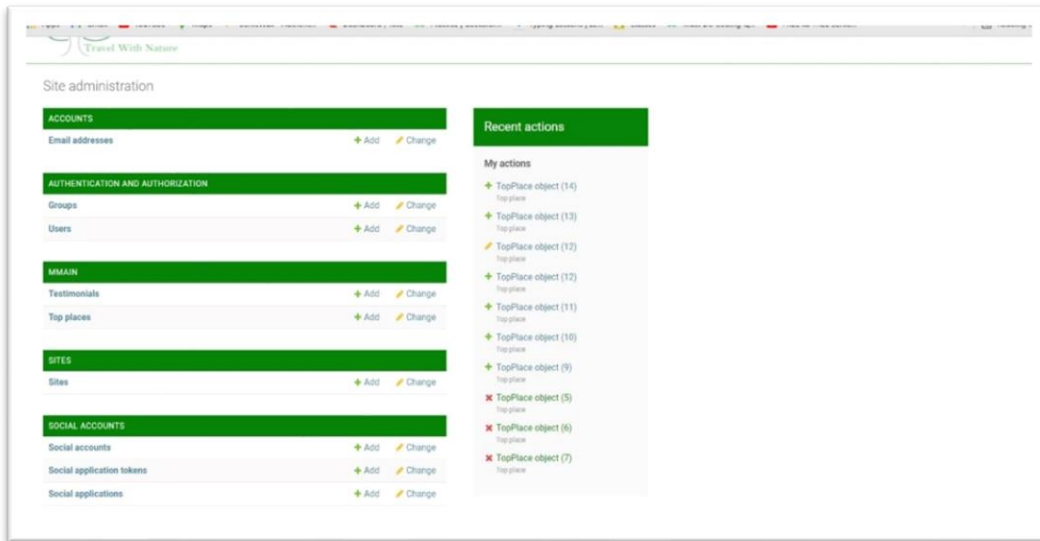
Fig(2)
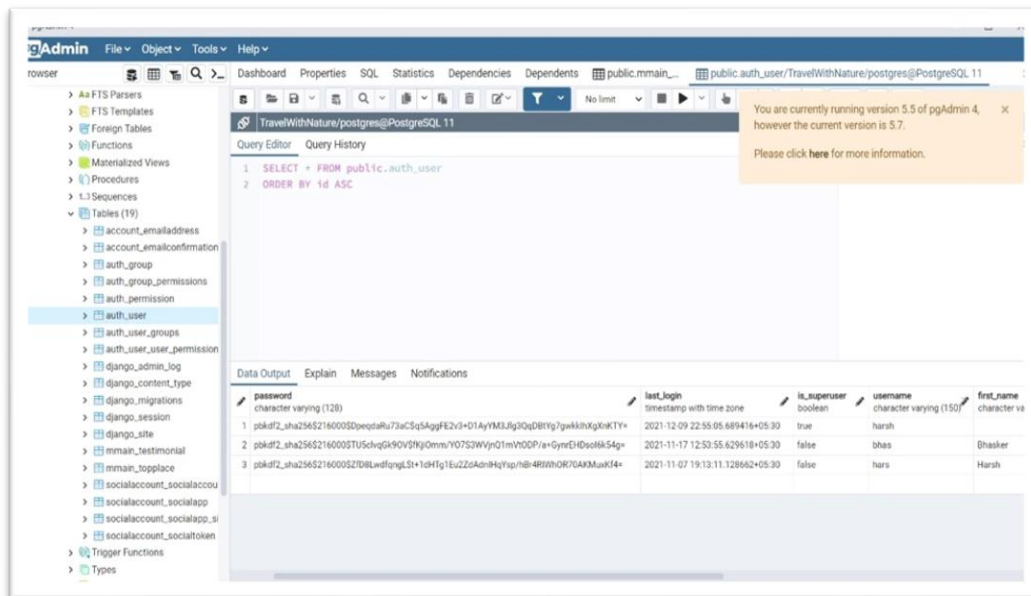
Fig(3)

3.3 COMMENT ANALYSIS

For Data processing, NLTK(Natural Language Toolkit) Library is used in our project.

NLTK is one of the NLP(Natural language Processing) libraries that help in statistical language processing and analyzing Data. It is used to Analysis the Comment section so that users can get to know whether the comment is negative, positive, or neutral and it will help customers for making better decisions about the place (fig-4).

The comment entered by the user is stored in a variable and then it is passed to the .py file which analysis the comment word by word and stores its polarity then on the basis of its polarity the comment will be categorized as positive, negative or neutral.
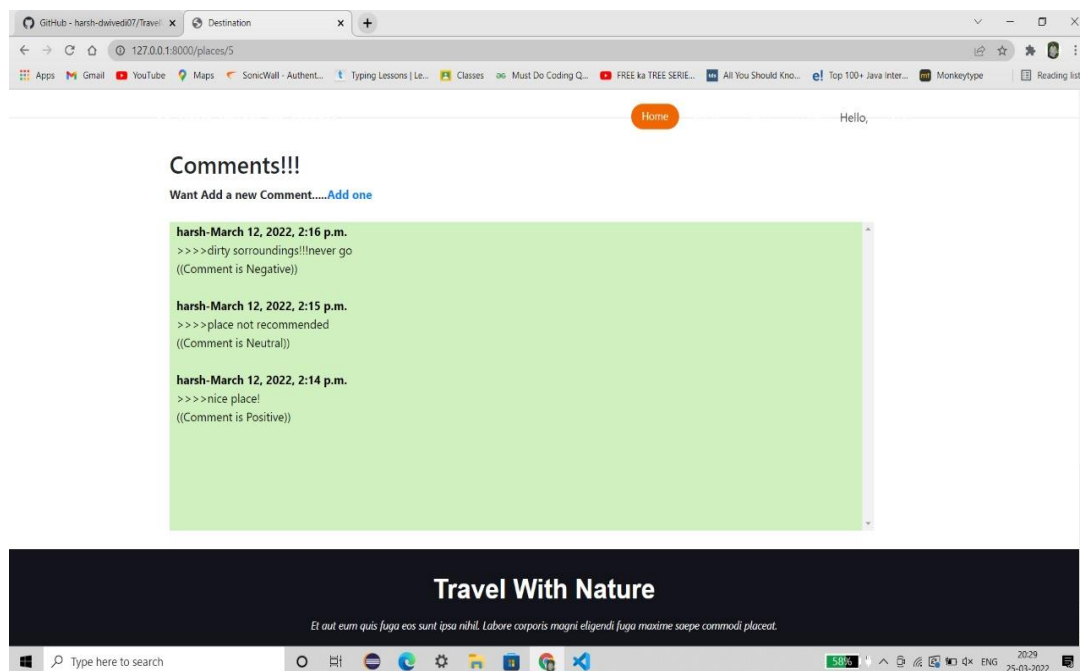
Fig(3)

4 CONCLUSION

A dynamic interactive website is created which contains information about the unexplored and unvisited region with the comment section area and comments are analyzed with the help of NLTK which helps in processing and analyzing data and deciding whether the comment is positive, negative, or neutral with the help of which travelers can decide whether they should visit that place or not, and tourism can help in overall development in those areas.

REFERENCES

1. Bhuiyan, H., Ara, J., Bardhan, R., & Islam, M. R. (2017, September). Retrieving YouTube video by sentiment analysis on user comment. In *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)* (pp. 474-478). IEEE.
2. Bhattacharjee, S., Das, A., Bhattacharya, U., Parui, S. K., & Roy, S. (2015, July). Sentiment analysis using cosine similarity measure. In *2015 IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS)* (pp. 27-32). IEEE.
3. Gupta, G., & Gupta, P. (2019, July). Twitter mining for sentiment analysis in tourism industry. In *2019 Third World Conference on Smart Trends in Systems Security and Sustainablity (WorldS4)* (pp. 302-306). IEEE.
4. Gao, S., Hao, J., & Fu, Y. (2015, June). The application and comparison of web services for sentiment analysis in tourism. In *2015 12th International Conference on Service Systems and Service Management (ICSSSM)* (pp. 1-6). IEEE.
5. Govindarajan, M. (2014). Sentiment analysis of restaurant reviews using hybrid classification method. *International Journal of Soft Computing and Artificial Intelligence*, *2*(1), 17-23.
6. Adeborna,E., & Siau, K.(2014). An approach to sentiment analysis-the case of airline quality rating.
7. Shi, H. X., & Li, X. J. (2011, July). A sentiment analysis model for hotel reviews based on supervised learning. In *2011 International Conference on Machine Learning and Cybernetics* (Vol. 3, pp. 950-954). IEEE.
8. Mutana, S. WEALTH CREATION THROUGH COMMUNITY-BASED CULTURAL TOURISM VENTURES: UNEXPLORED OPPORTUNITIES IN ZIMBABWEAN RURAL AREAS.
9. Ray, N., Dash, D. K., Sengupta, P. P., & Ghosh, S. (2012). Rural tourism and it's impact on socioeconomic condition: Evidence from West Bengal, India. *Global Journal of Business Research*, *6*(2), 11-22.
10. Aung, K. Z., & Myo, N. N. (2017, May). Sentiment analysis of students' comment using lexicon based approach. In *2017 IEEE/ACIS 16th international conference on computer and information science (ICIS)* (pp. 149-154). IEEE.