Research Paper

Classification of Machine learning approach for rating analysis for Cooking Recipes based Food Lovers

Dr .Ramesh T.Prajapati[1] Dr. Anil Suthar[2]

[1]HOD & Associate Prof. in CE/IT,Shree Swaminarayan Institute of Technology, Bhat, Gandhinagar, Gujarat, India

[2]Director, L.J. Institute of Engineering and Technology, Ahmedabad, Gujarat, India

Abstract

Online, you can find numerous recipes, some of which may be accurate while others may not. In this study, we present a data mining approach for identifying comments related to food recipes. We will employ a classification method to extract valuable insights from the data. By evaluating user reviews, the system will assign ratings to the recipes, making it easier for users to find reliable ones. This system proves beneficial for individuals seeking recipes online. Our plan involves training datasets comprising various food dishes, aiming to achieve precise and efficient detection of comments. This endeavor will contribute to the development of a robust application for food enthusiasts, addressing their concerns. In future it happens the real time system which is very helpful for foodies. Sentiment analysis in web embraces the problem of aggregating data in the web and extraction about opinions. Studying the opinions of customers helps to determine the people feeling about a product and how it is received in the market. Various commercial tools are available for sentiment analysis. In future in this system we will also proposed add new feature which users don't know.

Keywords - classification, support vector machine, comments, feedback from customer, classification, naïve bayes, cooking, sentiment analysis

1.Introduction

There is a diverse range of cuisines available, such as Indian, Mexican, Italian, and more, all created using similar ingredients. Traditional Indian cuisine encompasses a wide array of locally sourced spices, herbs, vegetables, and fruits. This system aims to assist food enthusiasts in discovering the best food dishes and recipes. It proves particularly useful for individuals who dine out at restaurants for lunch or dinner, as well as those who follow recipes on platforms like YouTube to cook at home. With the aid of a lens or masks, users can evaluate the quality of food using this system. It is also beneficial for those who watch recipes on YouTube but have not tasted the dishes themselves. They can employ the mask and lens to assess the food's quality. The system operates in real-time, allowing individuals dining at restaurants to capture a photo of their ordered dish and use the lens to evaluate its quality. Additionally, this system facilitates user ratings and comments, which benefit other people. By learning from a user's past preferences, the system can recommend new and untried recipes tailored to their taste.

Sentiment Analysis: it also referred to as opinion mining or emotion AI, involves the application of natural language processing, text analysis, computational linguistics, and biometrics to systematically detect, extract, measure, and analyze emotional states and subjective information. It encompasses the process of determining whether a piece of written content expresses a positive, negative, or neutral sentiment. Sentiment analysis serves as a valuable tool for data analysts in large organizations as it enables them to assess public opinion, conduct detailed market research, monitor brand and product reputation, and gain insights into customer experiences (17).

Classification: Classification is a crucial data mining technique that involves categorizing items in a collection into specific target categories or classes. The primary objective of classification is to make precise predictions about the target class for each case within the data. By utilizing classification models, it becomes possible to effectively determine the credit risk level of loan applicants, assigning them to categories such as low, medium, or high credit risks. This process enables accurate decision-making and risk assessment in various domains and scenarios (17).

2.Literature Review

In the first paper the author employed various tools to automatically extract comments on YouTube cooking recipes. In our system, the user starts by entering the specific recipe name on YouTube to search for the desired video. Using the YouTube Application Programming Interfaces provided by Google developers our system automatically retrieves the Uniform resource locator of the cooking recipe videos

Research Paper

related to the entered recipe. The tool then collects and stores the generated comments from these recipe videos in a database.

The next step involves feature selection and preprocessing of the dataset (8). Various algorithms are applied, and the system automatically filters opinion reviews from the collected comments. Texts that do not express any opinion are eliminated by classifying the comments into two classes, namely "opinion" and "other," using a Support Vector Machine classifier. The dataset is then preprocessed again, followed by feature selection.

Finally, the system classifies the filtered opinions into pre-defined classes, such as positive or negative, using the Support Vector Machine classifier. The system presents the results by counting and displaying the number of positive and negative comments. It calculates and showcases the percentage of each recipe video's comments and provides an overall percentage. Additionally, users of our system can access all the comments classified according to their polarity, allowing for a comprehensive view of the opinions expressed.

In the second research paper, Opinion Mining and Sentiment Analysis serve as crucial tools for gathering information and understanding people's thoughts. In our system, the user begins by entering the specific recipe name on YouTube to search for the desired video. Utilizing the YouTube APIs provided by "Google developers," our system automatically retrieves the URLs of the cooking recipe videos associated with the entered recipe (4). The system proceeds to collect and store the generated comments from these recipe videos in a database.

The next step involves splitting the comments into fragments based on previous steps. Then, the system extracts the polarity of these fragments by analyzing their features. Users can then obtain the results, which include the overall percentage and the number of Likes, Dislikes, and Views for each recipe video. These metrics provide valuable insights into the reception and popularity of the recipes among viewers.

In the third paper the author developed the Foodoholic application and conducted sentiment analysis on food recipes. The primary goal of the application was to rank various recipes based on reviews, considering the core ingredient used. This feature saves users' time by providing them with the best recipe options for a specific ingredient. Sentiment analysis is widely utilized to analyze reviews or social media content, allowing researchers to understand people's feelings and opinions on a particular topic.

The fundamental task of sentiment analysis is to analyze and classify the polarity of a given text, whether it is positive, negative, or neutral, within a document or sentence. The application further aims to identify opinions, feedback, or reviews and determine the sentiment being expressed. By conducting sentiment analysis, users can obtain valuable results regarding the overall sentiment of the reviews or feedback associated with the recipes in question.

The author said in fourth paper, a method is proposed for ranking recipes based on their "classicality" within a specified type of food. The process begins by removing inconsistencies present in the system, such as blank spaces, parentheses, special characters, and meaningless symbols. Next, the system focuses on ingredient recognition, ensuring accurate identification of the ingredients used in each recipe.

Afterward, algorithms are applied to the datasets, followed by the selection of relevant features. The recipes are then classified into different classes based on their characteristics (4). Finally, users are provided with the desired result, allowing them to access the ranked recipes based on their classicality within the specified food category.

In the fifth research paper, the authors introduce a tool designed to analyze the content within recipe comment texts. This tool serves to assist users in making informed decisions about food recipes. YouTube, being a popular platform, enables users to upload, rate, view, share, comment, and subscribe to videos and channels. The system described in the paper utilizes various machine learning algorithms, including K-Nearest Neighbor, Naïve Bayes, Support Vector Machine and Deep Learning, to achieve improved results. The first step involves text pre-processing, followed by text recognition by the system. In the subsequent step, machine learning algorithms and sentiment analysis are applied to the dataset (5).

As a result, users are provided with valuable insights and outcomes based on the reviews of the recipes. By leveraging these analytical techniques and algorithms, users can obtain more reliable information and make well-informed decisions about their food recipe choices.

3.Problem Statement

This study uses a dataset to predict the popularity of recipes, and it investigates several factors that can be extracted to enhance performance. The approach specifically suggests performing sentiment analysis on the reviews and treating the sentimental scores as a feature.

Proposed System

I proposed following algorithm. Steps are mentioned below.

1. Start

2. Input data from users in the form of Image, text and speech.

3. Convert text and speech into image.

4. Check Quality of recipe.

5. If it is good, then Generate the list of recipe or go to step 9

6. Check the ratings and feedback

7. Apply classification algorithm

8. Display or produce the list of recipe.

9. Stop.

In our proposed algorithm, users can provide input in the form of images, text, or speech. If the input is in text or speech form, it will be converted into an image. Subsequently, a lens or mask will assess the quality of the food. If the quality is deemed good, the system will generate a list of recipes. However, if the quality is not satisfactory, the system will halt. Users then have the option to review the ratings. If the ratings are positive, the system will apply a classification algorithm to generate the required recipe list. On the other hand, if the ratings are not favorable, the recipe will be removed from the list.

4.Results And Discussions

We taking dataset from Kaggle website which are containing thousands of datasets for different fields like Education, News, Politics etc. The shape of data is 255 rows and 9 columns.

Now, we convert the speech into text using Google speech recognition. For this we'll import speech recognition library. And if we've no connection of internet then the output will print to sorry, I did not get that otherwise they give us output which we'll speech.

Image quality is an open source software library for Automatic Image Quality Assessment. Here we use Brisque for check the quality of an image. Brisque (A)calculates the no-reference image quality score for image A using the Blind/Reference less Image Spatial Quality Evaluator (BRISQUE). brisque compare A to a default model computed from images of natural scenes with similar distortions. A smaller score indicates better perceptual quality. We check quality of burger image and we get 15.8674 burger image quality. We take the dataset reviews from Kaggle after that we check the number of ratings score which is in form of 1 to 5. And at the end we check the total number of ratings and plotting into a graph between score and total number of ratings.
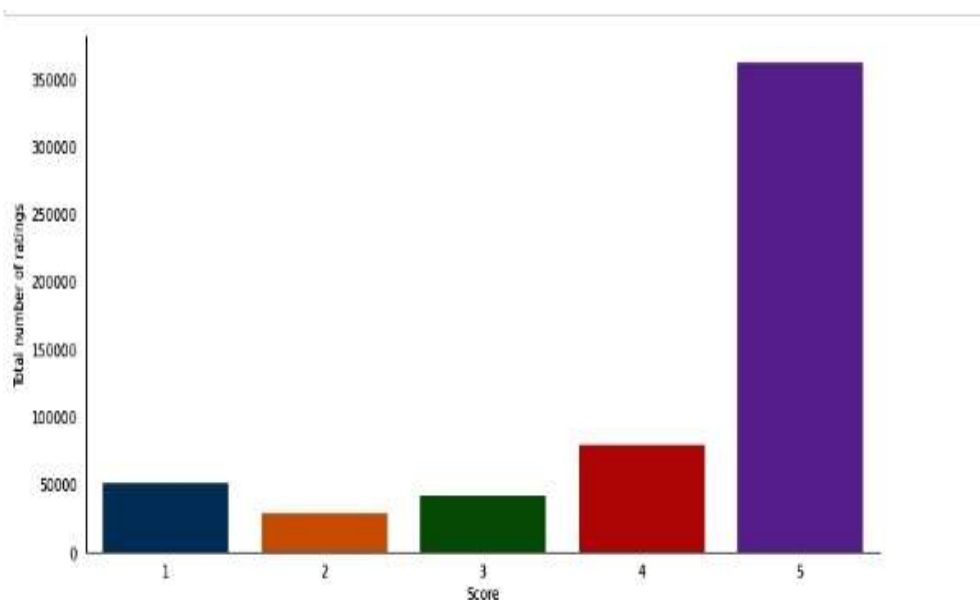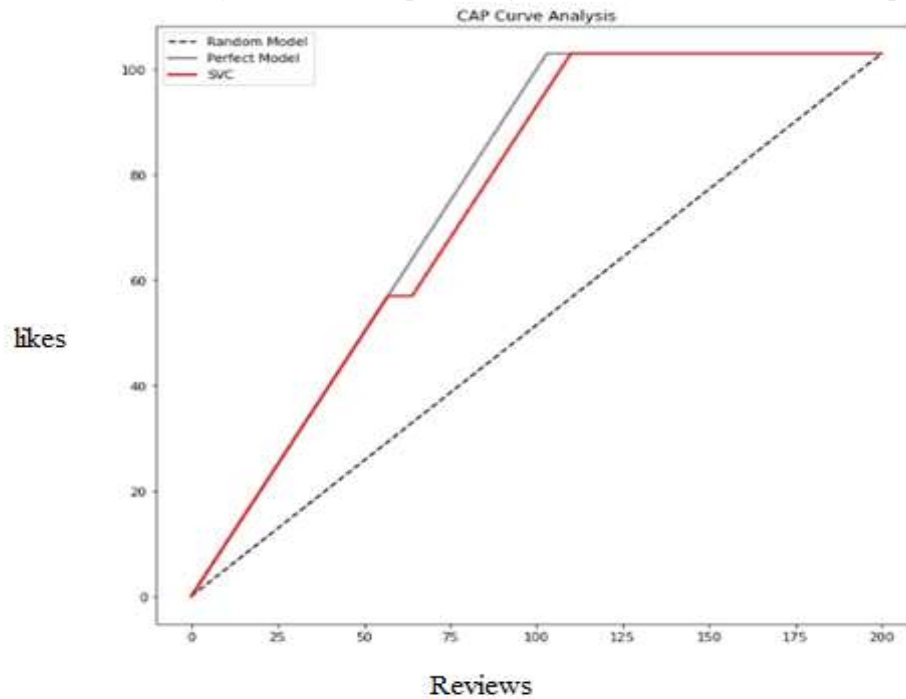
Research Paper

Fig 1- Ratings Graph

Here, we take total 256 recipes for check the ratings of recipes. And the graph indicates the total number of rating and score which the customers gave.

Fig. 2 Proposed Model Accuracy In the task of predicting whether the customer likes the product or not, based on reviews, the support vector

classifier has outperformed the Naïve Bayes. By analyzing the confusion matrices and CAP curves of each of the three classifiers, it is clear that the support vector classifier is best among all three. Here we get 86.2068% accuracy of our proposed model. Here we compare accuracy of base paper and existing system in pie chart.

Result Gathering	
Base Paper	**Existing System**
73.33%	86.21%

Table 1 Comparing Accuracy

Here we compare the accuracy in graph and we found accuracy of our proposed model is 86.21% and base paper accuracy is 73.33%.
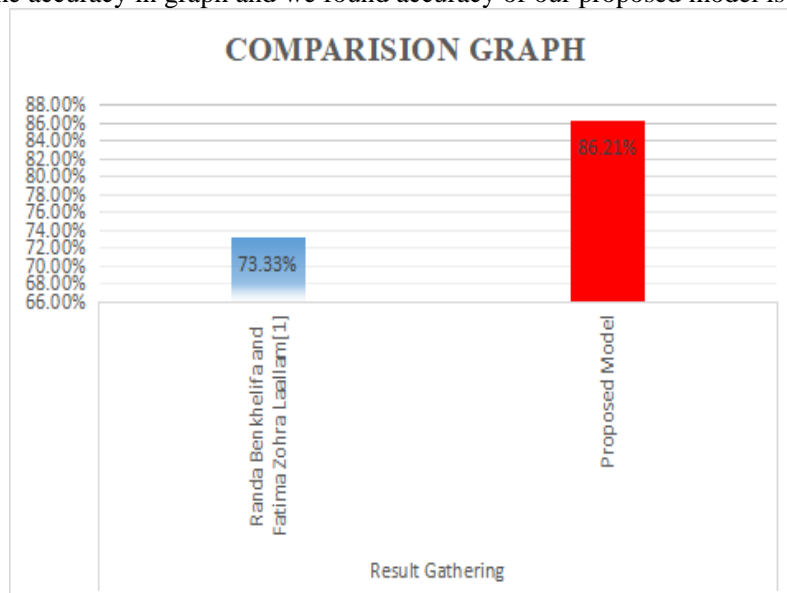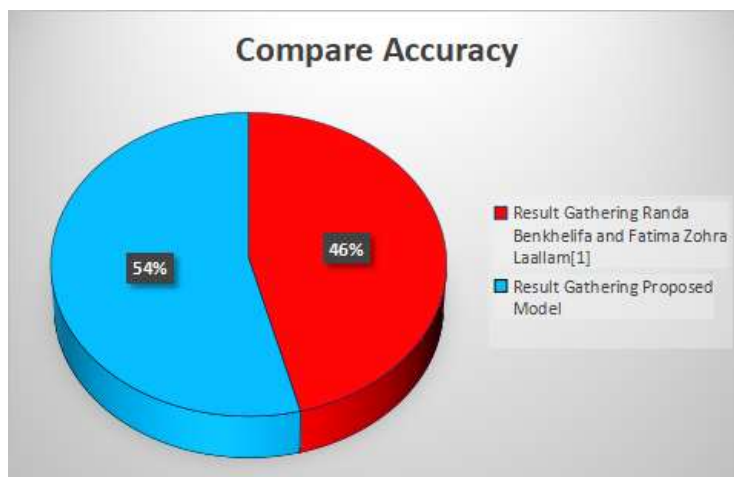
Research Paper

Fig. 3 Comparing

Accuracy using graph Here we compare the accuracy in pie chart and that's why we found accuracy of our proposed model is 54% and base paper accuracy is 46%.

Fig. 4 Comparing Accuracy using pie chart

5.Conclusion

This system serves as a valuable tool for assisting food enthusiasts and individuals in general. The meta-analysis conducted on recipes has provided motivation for further research, which can be applied in our specific environment. The proposed system can be evaluated and its efficiency determined through the use of specific parameters.

Opinions or comments expressed in various forms of content, including informational or knowledge-based content, hold significant value. Furthermore, opinions or comments from other individuals are highly beneficial in our decision-making processes. Therefore, the development of an automated technique capable of analyzing these opinions or comments becomes a valuable tool for assisting users, customers, consumers, and providers alike.The proposed system has produced a very good result using the score ratings. The limitation of this system is, it works better only for the open sentiments like rating or scores. The results were not promising for hidden sentiments. We proposed a methodology for mining the food reviews based on score combined with existing text analyzing packages.

6.Future Work

Our future work includes an investigation into more intelligent means of reasoning on food ratings when recipe ratings are known, and vice versa, on recipe ratings when food ratings are known. Also our main focus is to collect real time samples of food for individual restaurants and scan it with camera use of lens or mask. First training and after that testing of various food samples with standard dataset. And in future we build android or iOS based application for particular restaurants purpose also.

References

1. P. Chikersal, S. Poria, E. Cambria, A. Gelbukh, and C. E. Siong, "Modelling Public Sentiment in Twitter: Using Linguistic Patterns to Enhance Supervised Learning," in Computational Linguistics and Intelligent Text Processing: 16th International Conference, CICLing 2015, Cairo, Egypt, April 14-20, 2015, Proceedings, Part II, A. Gelbukh, Ed. Cham: Springer International Publishing, 2015, pp. 49–65 vol.10.
2. Pang, B.; Lee, L. A sentimental education: "Sentiment analysis using subjectivity summarization based on minimum cuts". In Proceedings of the 42nd Annual Meeting on Association for

Computational Linguistics, Barcelona, Spain, 21–26 July 2004; Association for Computational Linguistics: Stroudsburg, PA, USA, 2004; p. 271, MDPI.

3. "Association for Computational Linguistics (ACL 2004), Barcelona, ES," pp. 271–278 (2004) Bianchini, D., De Antonellis, V., De Franceschi, N., Melchiori, M.: PREFer: a prescriptionbased food recommender system. Comput. Stand. Interfaces, Springer vol. 54, 64–75 (2017)

4. Liu, C., Guo, C., Dakota, D., Rajagopalan, S., Li, W., Kübler, S.: "My curiosity was satisfied,but not in a GoodWay: predicting user ratings for online recipes. In: Proceedings of the Second Workshop on Natural Language Processing for Social Media (SocialNLP)", Dublin, Ireland,vol. 10, Springer, 24 August 2014, pp. 12–21 (2014)

5. Sebastiani, "A.: Machine learning in automated text categorization. ACM Comput. Surv"/Springer, vol. 34, 1–47 (2002)

6. Al-Smadi, M., Talafha, B., Al-Ayyoub, M., Jararweh, Y.: "Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews. Int. J. Mach. Learn." Cybern vol.10, Springer . 1–13 (2018)

7. Kamal, "A.: Subjectivity classification using machine learning techniques for mining feature opinion "pairs from web opinion sources"/Springer, vol.10 , New Delhi, India (2013)

8. Al-Smadi,M., Al-Ayyoub, M., Jararweh,Y.,Qawasmeh, O.: "Enhancing aspect-based sentiment analysis of Arabic hotels' reviews using morphological, syntactic and semantic features. Inf. Process. Manage" vol.10, Springer (2018)

9. Kaushik, A.; Naithani, S. "A Study on Sentiment Analysis: Methods and Tools. Int. J. Sci. Res. (IJSR)" 2014, 4, 287–292, MDPI. 17) Kaushik, A.; Naithani, S. "A comprehensive study of text mining approach. Int. J. Comput. Sci. Netw. Secur. (IJCSNS)" 2016, 16, 69, MDPI.

10. Benkhelifa, R., Laallam, F.Z.: "Facebook posts text classification to improve information filtering. In: Proceedings of the 12th International Conference on Web Information Systems and Technologies", pp. 202–207. Rome, Italy, vol.10, Springer (2016)

11. Chaturvedi, I., Ragusa, E., Gastaldo, P., Zunino, R., Cambria, E.: "Bayesian network based extreme learning machine for subjectivity detection. J. Franklin Inst." Vol.10 , Springer (2017)

12. Pang, A., Lee, L.:A sentimental education: "sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd Meeting of the Association for Computational Linguistics" (ACL 2004), Barcelona,vol.10, Springer, ES, pp. 271–278 (2004)

13. H. H. Wu, A. C. R. Tsai, R. T. H. Tsai and J. Y. j. Hsu, "Sentiment Value Propagation for an Integral Sentiment Dictionary Based on Commonsense Knowledge," 2011 International Conference on Technologies and Applications of Artificial Intelligence, Chung-Li, 2011, pp. 75-81, vol.10, Springer.

14. R. Tannahill," Food in history: Grown Trade Paperbacks",vol.10, IEEE 1989.

15. J. Flandrin, A. Sonnenfeld, and M. Montanari, "Food: a culinary history," Columbia University Press, vol.10, IEEE, 2013.

16. A. Nanetti, S. A. Cheong, and M. Filippov, "Interactive Global Histories:For a New Information Environment to Increase the Understanding of Historical Processes," In Proceedings of 2013 International Conference on Culture and Computing (Culture Computing), 2013, pp. 104-110, vol.10, IEEE.

17. Patel, Hetu and Suthar, Dr. Anilkumar, Sentiment Analysis for Rating Cooking Recipes Based on User's Feedback Using Data Mining Techniques (Classification) (January 23, 2021). GEDRAG & ORGANISATIE REVIEW, VOLUME 34 : ISSUE 01 – 2021.