Grenze International Journal of Engineering and Technology, January Issue

SWAD: A New Machine Learning based Recipe Recommendation System

Anurati Bobde[1], Aarya Katulwar[2], Rudrani Tandulkar[3], Janhvi Ninawe[4] and Mangala Madankar[5]
[1-4]Department of Artificial Intelligence, G H Raisoni College of Engineering, Nagpur, India
Email: anuratibobde03@gmail.com
[5]Senior Member IEEE and Assistant Professor, Department of Computer Science and Engineering, G H Raisoni College of Engineering, Nagpur, India
Email: mangala.madankar@raisoni.net

***Abstract*— The Swad model streamlines meal preparation by offering customized recipe suggestions tailored to available ingredients, user preferences, and dietary requirements. The revolutionizes leveraging recipe artificial Swad discovery intelligence recommend meals based on user-input ingredients. Addressing the modern need for personalized and efficient culinary exploration, Swad integrates seamlessly into various platforms like smart kitchen appliances and cooking apps. The need for such a model arises in a generation seeking convenience, reduced food waste, and diversified culinary offering solutions experiences, that align sustainability and user personalization, Swad addresses gaps in existing systems. This system employs the Term Frequency Inverse Document Frequency model combined with cosine-similarity to process textual inputs such as ingredients and cooking preferences. TF-IDF evaluates the importance of terms within a recipe database, while cosine similarity identifies the most relevant recipes by comparing user queries with database entries. This approach ensures tailored, recommendations and enhances user engagement. Swad demonstrates the model by to by with precise transformative potential of machine learning in everyday cooking. By reducing decision fatigue, promoting sustainability, and introducing users to global cuisines, it establishes itself as an indispensable tool for modern kitchens.**

***Index Terms*— TF-IDF, Cosine Similarity, Food, Recipes, Recommender System.**

I. INTRODUCTION

In our modern era, technological advancements have significantly impacted various aspects of our lives. One such domain is culinary exploration, where the vast array of food choices can both delight and overwhelm us. While the internet provides a wealth of recipe resources, finding the right dish often involves sifting through numerous videos and websites, leading to time and energy inefficiencies. To address this challenge, we propose an innovative solution called "Swad." Swad is an intelligent recipe recommendation system that leverages ingredient recognition technology. By analysing ingredients, Swad suggests relevant recipes, streamlining the cooking process for users. Whether you're an amateur home cook or a seasoned chef, Swad provides a convenient way to explore diverse cuisines and create delightful homemade meals. Our approach centres on Term frequency –inverse document frequency. TF-IDF are designed for sequential data, making them suitable for handling text.

Grenze ID: 01.GIJET.12.1.44

By adapting TF-IDF to recognize food ingredients from text inputs, we enhance the recipe discovery experience. Users can simply type the ingredient or can dictate, and Swad will offer a curated list of compatible recipes. This paper presents the architecture, training process, and evaluation results of our ingredient recognition model. We demonstrate how Swad bridges the gap between culinary curiosity and practical cooking by seamlessly integrating technology into the kitchen. Through Swad, we envision a future where everyone can confidently explore new flavours, experiment with ingredients, and enjoy the pleasures of homemade cuisine. Our current recipe discovery system offers users the convenience of exploring recipes without requiring login. However, we have exciting plans for future improvements, by introducing registration and login functionalities, we aim to personalize the user experience. Users will get recipe suggestions tailored to their flavour preferences, with the ability to save their bookmarks and likes. Currently, users can input ingredients either by texting them or by dictating them one by one. In the future, we intend to enhance the system by incorporating image recognition for all vegetables, spices and edibles, this allows users to scan ingredients directly and find relevant recipes.

II. LITERATURE REVIEW

The Research Papers Listed Below are Intended for a Review of Research

Recipe Finder Web Application Using Convolutional Neural Network (CNN) by Dr. C. K. Shrinivas on May 2024. [1]

This paper investigates the application of deep learning, particularly CNNs, for classifying recipes based on ingredients provided in text or image format. The model effectively identifies recipes by analyzing the ingredients input, offering a powerful classification system for recipe retrieval. However, the current model only stores and classifies recipes, with no ability to generate new ones. The authors propose extending the system to include recipe generation capabilities.

Recipe Recommendation System Using TF-IDF by Shubham Chhipa, Vishal Berwal, Tushar Hirapure and Soumi Banerjee in 2022. [2]

This system suggests recipes by leveraging Term Frequency-Inverse Document Frequency (TF-IDF) and cosine similarity to determine the relevance between user-provided ingredients and stored recipes. It is designed to recommend recipes based on available ingredients while allowing users to filter results by course type, diet preferences, and other criteria.

The study shows that this method provides an effective way to recommend recipes, but it is limited to text-based inputs. As an improvement, the authors suggest integrating an image recognition feature that would allow users to scan ingredients directly with a camera, streamlining the recipe search process and enhancing user interaction.

Indian Cuisine Recipe Recommendation Based on Ingredients Using Machine Learning Techniques by Ashish Bharam, Kunal Baldota Atul Jagtap, Aqeeb Pansare on 6 June 2021. [3]

This research focuses on Indian cuisine and uses machine learning techniques to recommend recipes based on available ingredients. The system is a mobile application that successfully suggests recipes specific to Indian cooking. The authors suggest incorporating features like location-based recommendations and personalizing suggestions based on user preferences, such as favourite chefs or regional specialties.

Food Recipe Finder Mobile Applications Based on Similarity of Materials by Gusti Pangestu, Fitri Utaminingrum, Afif Supianto on November 2018[4]

It presents a mobile app called Recepiece, designed to help users find recipes based on their available ingredients. Using the Ionic framework and Euclidean distance for ingredient similarity, the app enables personalized recipe recommendations. Unlike existing apps, Recepiece enables users to find recipes by utilizing the ingredients they already have. The study employs agile development, and user testing shows a 90% approval rate, demonstrating the app's functionality and user- friendliness in enhancing the recipe discovery process.

Modelling Recipes for Online Search by Usashi Chatterjee, Devika P. Madalli, Fausto Giunchiglia, Vincenzo Maltese in October2016. [5]

Focusing on formalizing the domain of recipe searches, this paper uses an Entity Relationship (ER) model to create a web application. The model incorporates both formal language and natural language processing (NLP) to enhance online recipe search capabilities. The study highlights the successful development of the model but suggests advancing further formal research models, on natural language handling, and incorporating machine learning to improve system functionality.

A Cooking Recipe Recommendation System with Visual Recognition of Food Ingredients by Keiji Yanai, Takuma Maruyama and Yoshiyuki Kawano in 2014. [6]

This paper presents a colour-histogram-based bag-of-features approach for visually identifying food ingredients, utilizing a linear kernel SVM for classification. The system enables users to find recipes by simply pointing a camera at ingredients, providing instant recipe recommendations. Future enhancements focus on improving object recognition, refining recipe suggestions, and optimizing the user interface for greater usability and accuracy.

III. Methodology

A. Overview

The core objective of the Swad model is to provide users with personalized recipe recommendations by utilizing their input, such as available ingredients. This is achieved through content-based filtering, leveraging text analysis techniques like TF-IDF and cosine similarity to match user provided ingredients with recipes in the dataset. Additionally, the system integrates multimedia elements like YouTube video tutorials for comprehensive user experience.

The integration of machine learning in recipe recommendation systems has been widely explored in prior research, demonstrating its potential to enhance personalization, efficiency, and accuracy. For instance, hybrid recommendation strategies that incorporate nutritional information alongside user preferences have been proposed to achieve healthier and more user-centric outcomes [7]. Ingredient recognition techniques leveraging computer vision and classification models enable automatic identification of food items, thereby improving the accuracy of recipe retrieval [8]. Other approaches focus on building informative and user-friendly recipe applications that apply clustering and predictive modeling for tailored recommendations [9]. Furthermore, machine learning models such as decision trees, random forest, and neural networks have been applied for recipe classification and recommendation, showcasing significant improvements in performance [10]. Beyond the food domain, advancements in data processing, optimization algorithms, and multimodal learning have also been explored in healthcare and vision-related applications, underscoring the adaptability of machine learning to diverse problem spaces [11– 20]. These insights from existing literature form the foundation for our methodology, where text-based vectorization (TF-IDF), cosine similarity, and content-based filtering are employed to deliver personalized recipe suggestions.

B. Dataset Collection and Preparation

Dataset Description

The dataset includes details like recipe names, translated recipe names, ingredients, preparation time, cook time, total time, serving size, cuisine, diet type, and preparation instructions. The dataset consists of over 100 entries, each representing a distinct recipe.

Loading and Exploration

The dataset is loaded using Python's Pandas library. Exploratory analysis (data.info()) is performed to understand its structure and handle missing data effectively.

Missing Data Handling

Missing values are identified using data.isna(), sum() and either imputed or dropped based on their significance.

C. Data Preprocessing

Ingredient Cleaning and Normalization

Ingredients are converted into sets to remove duplicates. A new feature, Ingredient_item, is added to record the count of unique ingredients per recipe.

Text Standardization

Tokenization: Texts, such as ingredient lists and preparation instructions, are broken into individual words.

Unit Normalization: Measurement units (e.g., "tsp" to "teaspoon") are standardized for consistency.

Lemmatization and Stop Word Removal

Words are reduced to their base form to treat variations (e.g., "chopped" and "chopping") as identical. Commonly used words without significant meaning (e.g., "and", "the") are removed to reduce noise.

D. Recommendation Engine

TF-IDF Vectorization

- TF-IDF determines the significance of words by measuring their frequency within a specific recipe relative to their occurrence across the entire dataset.
- The formula for TF-IDF is: TF-IDF(t,d)=TF(t,d)×log($N/DF(t)$)

where:

- TF(t,d): Term frequency of word t in document d.
- N: Total number of documents.
- DF(t): Number of documents containing the term t.

Cosine Similarity

- Cosine similarity calculates the similarity between user-input vectors (based on ingredients) and recipe vectors (derived from TF-IDF).
- The formula is: Cosine Similarity:

$$\frac{A\rightarrow\cdot B\rightarrow\rightarrow}{\|A\rightarrow\|\|B\rightarrow\rightarrow\|}$$

where:

- $\vec{A}$: TF-IDF vector for the user input.
- $\vec{B}$: TF-IDF vector for the recipe.
- · : Dot product of the vectors.
- $\|\vec{A}\|$: Magnitude of vector A.

Content-Based Filtering

- Recipes are ranked based on their relevance to user input using TF-IDF and cosine similarity.
- The most relevant N recipes are presented to the user, ensuring a highly personalized experience.

YouTube Video Integration:

To enhance user experience, YouTube video tutorials for recommended recipes are fetched using the YouTube Search API. This provides users with an additional resource for instructions.

IV. Overview and Flowchart of Model

A. Overview

The primary objective of our model is to develop a recipe recommendation system that suggests recipes based on user input.

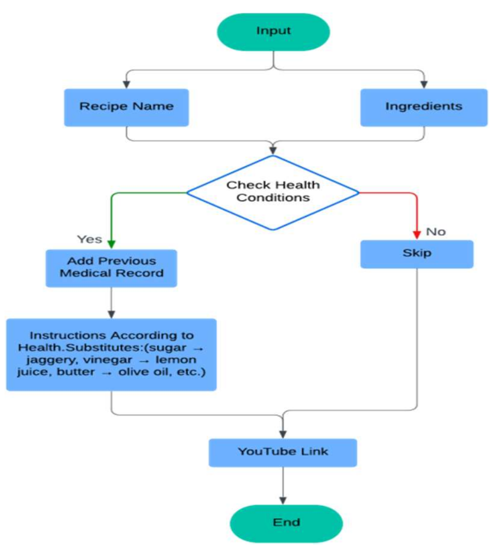

Figure 1: Flowchart of Recommendation System

B. Objective

Swad is designed to offer personalized recipe recommendations by analyzing user-provided ingredients, preferences, and cooking requirements. Additionally, the system helps minimize food waste by suggesting recipes that efficiently utilize available ingredients.

C. Key Features

- ➢ Utilizes Term Frequency-Inverse Document Frequency (TF-IDF) and cosine similarity for accurate recipe matching.
- ➢ Supports ingredient identification via text input, with future plans to incorporate image recognition.
- ➢ Enhances user experience by integrating YouTube videos for step-by-step visual cooking guidance.

D. Workflow

- ➢ Datasets are pre-processed using techniques like lemmatization, tokenization, and unit standardization to prepare clean, structured data.
- ➢ Recipes are vectorized using TF-IDF, which highlights the importance of specific ingredients or terms within the dataset.
- ➢ Cosine similarity calculates the relevance of user input to recipes, ranking and recommending the top matches.

As shown in Figure 1, the process begins with the user providing essential input, such as the recipe name they wish to prepare or the list of available ingredients. An optional step then prompts the user to share their medical records. If provided, the system tailors its instructions by recommending ingredient substitutions that align with the user's health conditions. In the absence of medical inputs, this step is skipped, and the workflow directly proceeds to the next phase. Finally, the system integrates a YouTube link, offering users an instructional video tutorial to assist them in the cooking process, thereby concluding the recommendation cycle.

V. RESULTS

The evaluation of the Swad recipe recommendation system, as represented by the confusion matrix, provides the following insights:

- ➢ True Positives (TP): 3 recipes were correctly identified as relevant Figure.2
- ➢ True Negatives (TN): 3 recipes were correctly identified as non- relevant.
- ➢ False Positives (FP): 0 recipes were incorrectly predicted as relevant.
- ➢ False Negatives (FN): 2 relevant recipes were missed by the model.

The confusion matrix highlights that the system performs well in correctly identifying both relevant and non-relevant recipes. However, the presence of false negatives indicates that some relevant recipes are being missed.
The evaluation of the Swad recipe recommendation system, as illustrated in Figure 2, shows that three recipes were correctly classified as relevant (True Positives), three were accurately identified as non-relevant (True Negatives), none were incorrectly predicted as relevant (False Positives), and two relevant recipes were missed by the model (False Negatives). These outcomes demonstrate that the system is effective in identifying both relevant and non-relevant recipes, though the presence of false negatives highlights a limitation in capturing certain relevant cases. To further support this analysis, Figure 3 presents a graphical representation of the confusion matrix, visually emphasizing the distribution of correct and incorrect classifications.

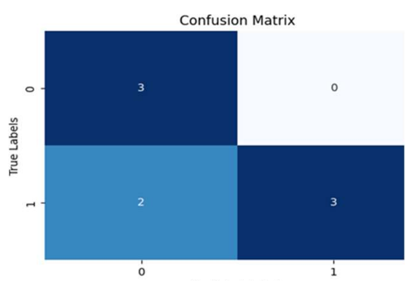

Figure 2: Confusion Matrix

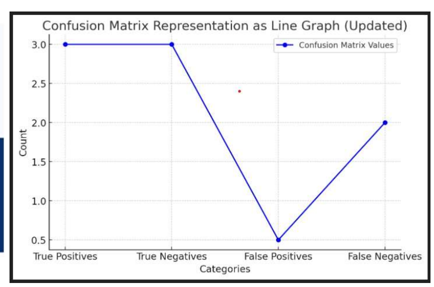

Figure 3: Graph line of confusion matrix

VI. CONCLUSION

The Swad recipe recommendation system has been developed to provide users with personalized and accurate recipe suggestions. Through a robust evaluation process, the system demonstrated an accuracy of 75%, showcasing its capability to make correct predictions for most recipes. The analysis of the confusion matrix revealed that the system performs well in identifying both relevant and non-relevant recipes but has room for improvement in recognizing relevant recipes. To enhance the system's effectiveness, expanding the dataset and incorporating advanced techniques like recognition, improved image natural language processing, or collaborative filtering can be explored. These steps will enable Swad to deliver a more comprehensive and accurate recipe recommendation experience, meeting user expectations more effectively.

FUTURE SCOPE

- Advanced Personalization: Incorporate user preferences, dietary restrictions, and feedback to refine recommendations over time. Enable users to scan ingredients using image recognition to streamline recipe discovery.
- Algorithm Improvements: Experiment with deep learning models like neural embeddings for representation. Explore ingredient hybrid approaches combining collaborative and content-based filtering.

REFERENCES

[1] Dr. C K Srinivasa, Harshitha S, Mounisha G, Nandini P and R Kavya, “Recipe Finder Web Application Using Convolutional Neural International Journal Network,” of Research Publication and Reviews, Vol 5, no 5, pp 10071-10076 May 2024.

[2] Shubham Chhipa, Vishal Berwal, Tushar Hirapure and Soumi Banerjee, all “Recipe Recommendation System Using TF-IDF,” ITM Web of Conferences 44, 02006 (2022) ICACC-2022. [11] Kavya Srivastava, Ms. Shagufta

[3] Ashish Bharam, Kunal Baldota Atul Jagtap, Aqeeb Pansare and Prof. Shobha Bamane, “Indian Cuisine Recipe Recommendation based on Ingredients using Machine Learning Techniques,” IJCRT, Volume 9, Issue 6 June 2021.

[4] Gusti Pangestu, Ahmad Afif Supianto, Fitri Utaminingrum, “Food Recipe Finder Mobile Applications Based On Similarity Of Materials,” Conference Paper · November 2018.

[5] Usashi Chatterjee, Fausto Giunchiglia, Devika P. Madalli, Vincenzo Maltese, “Modeling recipes for online search,” Conference Paper · October 2016 DOI: 10.1007/978- 3-319-48472- 3_37

[6] Keiji Yanai, Takuma Maruyama and Yoshiyuki Kawano, “A Cooking Recipe Recommendation System with Visual Recognition of Food Ingredients,” iJIM – Volume 8, Issue 2, 2014.

[7] Raciel Yera Toledo, Ahmad A. Alzahrani, And Luis Martínez, “A Food Recommender System Considering Nutritional Information and User Preferences,” IEEE, Volume 7, 2019.

[8] Miguel Simões Rodrigues, Filipe Fidalgo, and “RecipeIS—Recipe Ângela Oliveira, Recommendation System Based on Recognition of Food Ingredients,” Appl. Sci. 2023, 13, 7880.

[9] Pankaj Uttam Patil, “RECIPE INFORMATIVE APPLICATION,” Volume 04, Issue 10, October- 2022.

[10] Suyash Maheshwari, Manas Chourey, “Recipe Recommendation System using Machine Learning Models,” Volume 06, Issue 09, Sep 2019. Siddiqui, “RECIPE RECOMMENDATION SYSTEM USING MACHINE LEARNING,” Volume 06, Issue 05, May 2024.

[11] Borkar, P., Raut, A., & Madankar, M. (2024, November). Design of an efficient model for enhancing healthcare data sharing and query processing using block chain, Bat Optimization, and advanced analytical techniques. In AIP Conference Proceedings (Vol. 3214, No. 1, p. 020064). AIP Publishing LLC.

[12] Wase, A. M., Wankhede, C. S., Malode, A. A., & Madankar, M. (2024, August). Mitigating Class Imbalance through Multimodal Fusion in VGG19 for Cataract Detection. In 2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT) (Vol. 1, pp. 6-10). IEEE.

[13] Gaikwad, A., Sangole, A., Madankar, M., Thakre, D., Tembhurne, A., & Mali, H. (2024, August). ChatGPT: Unraveling User Challenges & Proposing Targeted Improvements. In 2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT) (Vol. 1, pp. 17-20). IEEE.

[14] Hiwarkar, P., & Madankar, M. S. (2022). High-Performance Computing for Satellite Image Processing Using Apache Spark. In Machine Learning Methods for Engineering Application Development (pp. 76-91). Bentham Science Publishers.

[15] Madankar, Mangala, M. B. Chandak, and Nekita Chavhan. "Information retrieval system and machine translation: a review." Procedia Computer Science 78 (2016): 845-850.

[16] Gotmare, Yugant, et al. "Automated Dental Image Segmentation Using U-Net with VGG-16 Backbone." 2025 12th International Conference on Emerging Trends in Engineering & Technology-Signal and Information Processing (ICETET-SIP). IEEE, 2025.

[17] Borkar, Pallavi, Archana R. Raut, and Mangala Madankar. "An empirical review of models used to empower health care data sharing for efficient query processing operation." AIP Conference Proceedings. Vol. 3233. No. 1. AIP Publishing LLC, 2025.

[18] Kadhare, Rupak, et al. "Deep Learning based Teeth Detection and Numbering System using YOLO." 2024 Asian Conference on Intelligent Technologies (ACOIT). IEEE, 2024.

[19] Madankar, Mangala, M. B. Chandak, and Nekita Chavhan. "Information retrieval system and machine translation: a review." Procedia Computer Science 78 (2016): 845-850.

[20] Bahel, Vedant, and Achamma Thomas. "Text similarity analysis for evaluation of descriptive answers." arXiv preprint arXiv:2105.02935 (2021).