

# Discovery of Recipes Based on Ingredients using Machine Learning

S. Praveen<sup>1</sup>, M.V. Prithivi Raj<sup>2</sup>, R. Poovarasan<sup>3</sup>, V. Thiruvankadam<sup>4</sup>, M. Kavinkumar<sup>5</sup>

<sup>1,2,3,4</sup>B.Tech, Dept. Of Information Technology, Valliammai Engineering College, Chennai, Tamil Nadu

<sup>5</sup>Asst. Prof, Dept. Of Information Technology, Valliammai Engineering College, Chennai, Tamil Nadu

\*\*\*

**Abstract** - There are millions of recipes which are uploaded on websites. But still, the user cannot find what dishes can be cooked by using the ingredients the user has. A deep learning approach is focused on finding out the cuisine involved in a certain recipe through various ingredients used in it. Some ingredients have a unique feature which will be available only at a certain part of the country. A novel method is used to find out the best-suggested recipe through some available ingredients based on use. A common dataset is created which consists of a global recipe and their procedure to cook it. A recipe is suggested as well as a web link to view the recipe procedure. A recommendation method is used in which the ingredients available by the user is taken as input and analyzation process is done with the help of data-set collected, and the appropriate dishes or recipes is recommended to the user by Machine Learning using K-Nearest Neighbors algorithm. The discovered output is visualized in the form of Website which is more convenient and user-friendly.

**Key Words:** Ingredients, Recipes, Machine Learning, K-Nearest Neighbors Algorithm.

## 1. INTRODUCTION

In olden days people use books which explain to manage the household and prepare food. But nowadays people try new dishes and experiment with food by using popular recipe sharing sites, and cooking websites. Here numerous amounts of recipes are found along with the ingredients and cooking methods. The internet enables people to share knowledge. Food recipe sharing platforms have been prevalent and receive their due attention from web surfers as well as food enthusiasts. Food is an inseparable part of our lives. It has been observed that ingredients and recipes are often considered when an individual chooses to eat. Influenced by ingredients and style of cooking, a cuisine can have several hundred or thousands of recipes for different dishes. The user currently cooks a dish with reference to recipes on various types of information source such as book and website. Though many books for cooking has been published, most people recently use recipes on Website such as "Foodnetwork, Jamie Oliver, FoodClub". Cooking beginners are encouraged by the convenience to access so many ingredients and varied recipes. At first, instance when a user needs to cook a recipe, the recipe name is surfed on the website. A recipe on Website shows the ingredients that are needed for a dish and the procedure of the cooking. But the problem is, the user cannot identify what are the dishes can be cooked by using the ingredients available by the user. To

overcome these problems, Machine Learning approach is used which enables to suggest the recipes based on the available ingredients by the user. This way of searching makes the user, the selection of recipes in a smarter way, and makes the household food maker easier. The objective is to reduce the selection of recipes in an easier way by using ingredients as input. The main problem is the user cannot find the recipes by using the available ingredients. It mainly focuses on finding out the cuisine involved in a certain recipe through various ingredients used in it. Some ingredients have a unique feature which will be available only at a certain part of the country. A novel method is used to find out the best-suggested recipe through some of the available ingredients with the user. A common dataset is created which consists of a global recipe and their procedure to cook it. The recipe is suggested as well as a web link to see the recipe procedure is provided. The database is created by the scrapping of data from the food websites and data is segregated in the .xls or CSV file format. The obtained data is not in the proper format and it is preprocessed and converted into a usable form. The ingredients available by the user is taken as an input using a web search bar and it is compared with a database which is followed by recommendation of the recipe based on the machine learning KNN algorithm.

## 2. RELATED WORKS

A deep learning approach is focused on finding out the cuisine involved in a certain recipe through various ingredients used in it. Some ingredients have a unique feature which will be available only at a certain part of the country. In the analyzation, it is defined that there are 12 major cuisines in the world and also many more. Support vector machine and random forest algorithm are used for the classification of cuisine. The linkage between the recipe ingredients and the identification of the cuisine is done. The problem can be tackled by machine learning algorithms like SVM, neural networks and Naïve Bayes. A framework is given for the typicality analysis of the combination of the ingredients. The list of ingredients can be arranged by adjusting the typicality value by adding or removing ingredients iteratively. The cooking beginners are encouraged by the convenience to access so many and varied recipes. The recipes on websites show the ingredients, however, some of the listed ingredients cannot be used for the cooking. A novel method is designed to find out the best-suggested recipe through some of the available ingredients with us. Here we create a common dataset which consists of

a global recipe and their procedure to cook it. In the proposed system we find out the suggested recipe as well as we show you a web link to see the recipe procedure. The database is created by scraping technique where recipe name, ingredients and the procedure link is collected from websites like *food network, Jami Oliver*. The collected data is stored in an excel file. The recipe matching algorithm is used for selection of related recipes for the given ingredients and in the advanced recommendation, the machine learning KNN algorithm is used for classification purpose.

### 3. METHODOLOGY

#### 3.1 System Architecture Diagram

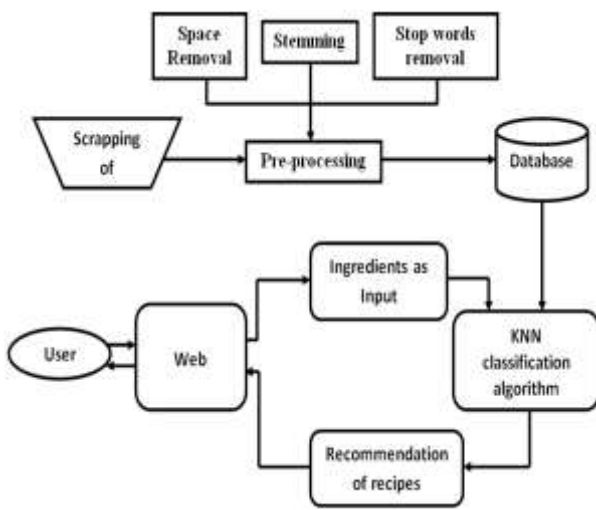


Fig -1: System Architecture Diagram

Fig -1 is the Overall architecture diagram of the Discovery of the recipe based on ingredients using machine learning. In the beginning, the database is created by scraping of particular data like Recipe name, ingredients and web link for the recipe's procedure from familiar food websites like "*foodnetwork, Jami Oliver, food club*". In general, the scraped data is not in the proper format and it is converted into a usable form by removing space and stop words. The preprocessed data is stored in the .xls file or .CSV file format. The available ingredients by the user are taken as input through the web interface. Using KNN Classification algorithm the input text is converted into a vector format and classified with the trained data and appropriate class is defined where the matched recipes are recommended to the user. The final output is visualized through websites which is user friendly.

#### 3.2 Data Preprocessing

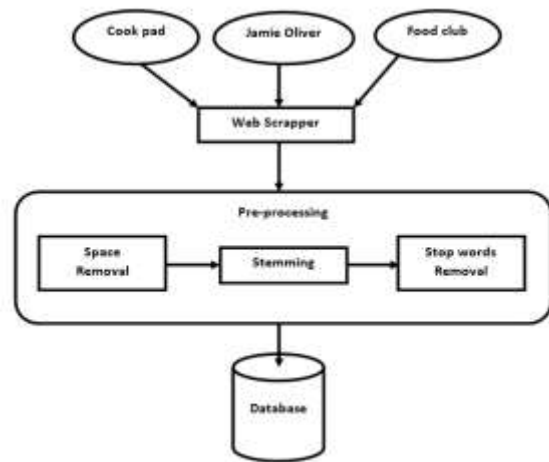


Fig -2: Data Creation & Preprocessing

Fig -2 represents the preprocessing procedures. Crawling usually refers to dealing with huge amount of data sets where you develop your own crawlers which crawl to deepest of the web pages. Data scraping on the other side refers to retrieving information or data from any sources. A Web crawler is an internet bot which systematically browses the World Wide Web typically for the purpose of web indexing. Web search engines and some other sites use web crawling or spidering software to update their web contents or indexes of other sites web content. With Chrome extensions like Agent Web Scraper, Data Miner, Spider web scraper the data is scraped from the web pages like cookpad.com, russianfood.com, thekitchen.com. Usually, the crawled data is in an unstructured format. So the data cannot be directly stored in the database. Hence the data is cleaned and converted into a structured format. The preprocessed data is stored in the database in the .xls file or .CSV file format.

#### 3.3 KNN Classification

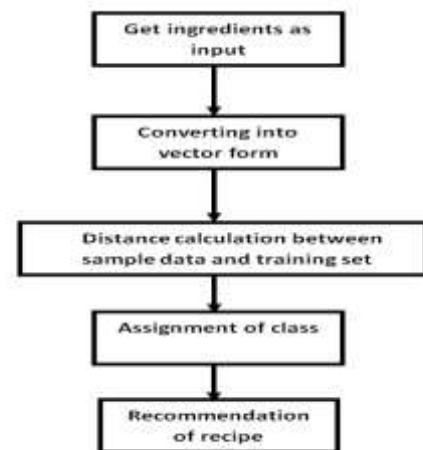


Fig -3: KNN Classification

Fig-3 represents that the KNN classification algorithm is divided into three parts namely Sampling, Training, and Class. In sampling, the ingredients are taken as input in the text format and converted into vector form. In training, the predefined vector values are assigned to the ingredients. Based on the weightage of ingredients of the recipes the class is assigned. The nearest distance between the sampling and training data is compared and the obtained value is assigned to the appropriate class. The assigned class recipes are recommended to the user.

### 3.4 User Interface



Fig -4: User Interface

Figure No. 4 Represents When the user access the website for the first time the registration is done by filling the basic details and login username and password is created. There will be three text box the user can type the three type of ingredients that user have while clicking the find button the appropriate recipes will be recommended the user through web links by clicking the link it will traverse to the food websites.

### 4. RESULT

The data scrapped from the website is prone to noise where there are lots of typos, stop word and unwanted line spacing. In order to clean the noise, preprocessing is done and obtained data is stored in the database. With the help of KNN classification algorithm, the input ingredients are processed and the suggested class is recommended as a recipe which is visualized in the website.

### 5. CONCLUSION

This system will help to select the recipe by using the available ingredients by the user. The selection of the recipe is in a smarter way and user-friendly. Various recipes can be classified according to the ingredients. In future enhancement, User can give input also in image format where the Image Processing is applied in detection of Ingredients. With the advanced machine learning techniques, the discovery of new recipe using available ingredients by the user can be applied in a smarter way

### REFERENCES

- [1] Li, Y., Meng, X., Wang, L., Li, Q. Recipecrawler: Collecting recipe data from www incrementally. In: *Proc. of the International Conference Web-Age Information Management. 2006, p. 263-274.*
- [2] M. Ueda, M. Takahata, and S. Nakajima, "User's food preference extraction for personalized cooking recipe recommendation," *Proc. 2nd Workshop on Semantic Personalized Information Management: Retrieval and Recommendation*, pp. 98-105, Oct. 2011.
- [3] 19th International Conference on Knowledge based and Intelligent Information and Engineering Systems Alternative-ingredient Recommendation Based on Co-occurrence Relation on Recipe Database
- [4] Melville, Prem, Raymond J. Mooney, and Ramadass Nagarajan. "Content-boosted collaborative filtering for improved recommendation" In *Aaai/iaai*, pp. 187-192. 2002.
- [5] Using *k*NN Model-based Approach for Automatic Text Categorization Gongde Guo<sup>1</sup>, Hui Wang<sup>1</sup>, David Bell<sup>2</sup>, Yaxin Bi<sup>2</sup>, and Kieran Greer<sup>1</sup> School of Computing and Mathematics, University of Ulster Newtownabbey, BT37 0QB, Northern Ireland, UK 1 {G.Guo, H.Wang, Krc.Greer@ulst.ac.uk} School of Computer Science, Queen's University Belfast Belfast, BT7 1NN, UK 2 {DA.Bell, Y.Bi}@qub.ac.uk

### BIOGRAPHIES



Mr. Praveen.S is studying in IV year of B.Tech (Information Technology) at Valliammai Engineering College, SRM Nagar Kattankulathur, Kancheepuram. (Tamil Nadu). His areas of interest are Data Science, Big Data processing etc.



Mr. Prithivi Raj.M.V is studying in IV year of B.Tech (Information Technology) at Valliammai Engineering College, SRM Nagar Kattankulathur, Kancheepuram. (Tamil Nadu). His areas of interest are Data Analytics, quantum computing etc.



Mr. Poovarasana.R is studying in IV year of B.Tech (Information Technology) at Valliammai Engineering College, SRM Nagar Kattankulathur, Kancheepuram. (Tamil Nadu). His areas of interest are Web Design and development, Networking.



Mr. Thiruvenkadam.V is studying in IV year of B.Tech (Information Technology) at Valliammai Engineering College, SRM Nagar Kattankulathur, Kancheepuram. (Tamil Nadu). His areas of interest are Web Design, Networking etc.



Mr.M.Kavinkumar.is working as an Assistant. Professor, Department of Information Technology at Valliammai Engineering College, SRM Nagar Kattankulathur, Kancheepuram. (Tamil Nadu). He has 7 years experience in Teaching and his research area is Data Mining, Knowledge Engineering and computational Linguistics.