Machine Learning Model for Predicting the Cuisine Category from a Dish Ingredients

Faisal Alshanketi
Department of Computer Science
Jazan University
Jazan, KSA
falshanketi@jazanu.edu.sa

2023 International Conference on Smart Computing and Application (ICSCA) | 979-8-3503-4705-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICSCA57840.2023.10087436

***Abstract*—Since ancient times, discussing and recording cooking recipes has been a widespread tradition. The ensuing massive collection of recipes and ingredients has the potential to advance significantly our understanding of both the basics of cooking and food pairing. Many sites have developed cooking suggestion systems or recipe engines in response to the rising popularity of food-based and recipe-sharing. Although this recommendation system provides dish suggestions, it cannot take advantage of the relationship between ingredients and cuisines. Exploring different cultures are achieved by tasting their food. Each country has its own traditional and unique cuisine. The cuisine can be determined by looking at the dish's ingredients. Therefore, the relationship between components and cuisines to predict the recipe cuisine is a motivating task to explore. The problem can be considered a classification problem with several classes. This research project attempts to shift focus away from recipe recommendations and toward investigating and examining the fundamental relationship between the components used in recipes and the various cultures. Support vector machine and associative classification, two basic classification approaches in data science, were used to examine the relationship between different recipes and their ingredient sets. On the dataset sourced from Kaggle, the tests were run. It gave a thorough understanding of other cuisines, ingredient patterns, and the key components of a good recipe. Additionally, the effectiveness of the classifiers used to forecast the cuisines was compared.**

***Index Terms*—cuisine prediction, recipe ingredients, machine learning**

I. Introduction

Because there are so many sites for sharing recipes, it is now easier to find online cooking recipes with lists of ingredients and step-by-step instructions. The creation of a recipe suggestion and retrieval system is something that many sites that share recipes have been working on. The recommendations made by these systems take into consideration the user's preferences [1], [2], the requested ingredients [3], the recipe information [4], and the cooking steps [5].

Users' preferences about recipe cuisines are thought to be one of the primary factors in determining what they decide to order. Ingredients are the second most crucial factor in choosing a cuisine's taste. If the cuisine associated with a dish of interest can be defined automatically, the quality of the recipe suggestions could be improved. To our knowledge, few papers have examined the combinations of ingredients and the recipe cuisines. As a result, the first thing that has been done in this study is to make an effort to analyze the fundamental relationship between cooking and ingredients.

The classification techniques are used to investigate the ingredient pairings in different cuisines. This paper aims to predict the cuisines of recipes by using recipes as examples, ingredients as features, and cuisines as the target label.

By assessing classification accuracy, we may see which cuisines have comparable components and are thus classed together. The classification standards evaluate each cuisine's elements by comparing ten machine learning algorithms to achieve the objectives and measuring the system performance using confusion matrices. Also, providing classification rules for each cuisine's components are measurements for predicting model performance.

Classifying a recipe cuisine may lead to four advantages and uses. First, although most recipes do not include cuisine information, labeling may help consumers make judgments. Second, cuisine information might be used to filter recipe recommendations. Third, dishes from the same cuisine should be combined while arranging a dinner. Fourth, identify a dish's cuisine.

The paper is organized as follows. Section II summarizes and discusses related work on the cuisine ingredients category, and then Section III presents the dataset used in our study. In Section IV, we describe the experimental methodology of the proposed approach. Section V presents our method for predicting the cuisine category from a dish's ingredients. In Section VI, we make concluding remarks and summarize future work.

II. Related Work

Several researchers have conducted studies on the analysis of food components, including their fundamental constituents and the primary flavors. These studies were conducted in particular to understand the combination as well as food matching for the food and hospitality businesses. The research was conducted on a wide range of sizes and levels; some of the investigations looked into the fundamental chemicals, while others looked into the flavor network.

The process of recipe retrieval and recommendation has been the focus of study relevant to cooking. Kalas is a social navigation system for culinary recipes. The system was created by Svensson et al. [6] and is considered one of the earliest systems. Ueda et al. proposed a system for providing individualized recipe recommendations based on the users' preferred foods [1], [2]. The user's history of looking up recipes and planning meals is used to determine what ingredients they like in their food.

Ahn et al. [1] create a ”flavor network” using a dataset of approximately 56,000 recipes and relate these recipes to the geographical groupings of countries in order to examine the culturally specific ingredient relationships that exist between dishes.

A user's food choice, expressed in terms of the ingredients used to prepare it, may be deduced from the user's history of recipe browsing and menu planning activities. A hybrid semantic item model was introduced by Xie et al. [3] for the purpose of recipe search by example. The hybrid semantic item model is used to describe the many various types of characteristics that may be found in recipe data.

Forbes introduces a technique for the selection of recipes that combines recipe material into the matrix factorization algorithm [4]. The experiment results indicated that the algorithm not only enhances the accuracy of the suggestion, but it is also helpful for switching out components and developing other versions of recipes.

Wang et al. [5] conducted an experimental investigation that concentrated on Chinese recipes represented as directed graphs. They offered a substructure similarity evaluation based on the frequent graph mining. The majority of research models recipes in terms of the ingredients that are used, while Wang et al. [5] used a different approach. To our knowledge, very little work is put into evaluating the association between the cuisine of a dish and the ingredients, despite the fact that there have been a number of effective ideas for the recommendation of recipes. Because of reason, the purpose of this research is to make the first effort to explore whether or not the cuisines of recipes may be determined by using the components of recipes.

Su et al. [7] use machine learning to classify cuisines and ingredients and predict cuisines from recipes. Like ours, it's based on a large data set of recipes (226K from food.com). In terms of prediction accuracy and F-measure, they achieved 75% using SVM.

Jayaraman et al. [8] analyzed the association between recipe cuisines and ingredient lists. The tests were run on a dataset that was compiled from a diverse range of sources, such as Food.com, Epicurious, and Yummly where they achieved accuracy scores for NB, RF, SVM, and Multinominal Regression were 73.5%, 76.2%, 79, and 78.8% respectively.

Automatic recipe cuisine categorization was carried out by Kalajdziski et al. [9] using text preprocessing and feature selection methods (such TF-IDF and Bag of Words techniques). Following the experiments, the optimum combination of feature selection and classifier was chosen to be the resulting system. The comparison between the NB classifier, an artificial neural network, and SVM algorithms produced classification accuracy of 74.8%, 79.1%, and 80.9% respectively.

III. Dataset

In our experimental analysis, the public dataset containing 39774 recipes associated with a particular cuisine and a particular set of ingredients was used. It is obtained from the Kaggle website and made available online [1] as a JSON file containing an array of objects, each listing id, cuisines, and ingredients. The discovery of the first analysis revealed that 6714 unique ingredients represent 20 cuisines. A sample JSON file shows the ingredients object as an array. Table I shows ten samples of the recipe id, the type of cuisine, and each recipe's ingredients list.

TABLE I
Dataset Description

id	cuisine	ingredients
10259	greek	['romaine lettuce', 'black olives', 'grape tomatoes', 'garlic', 'pepper', 'purple onion', 'seasoning', 'garbanzo beans', 'feta cheese crumbles']
25693	southern_us	['plain flour', 'ground pepper', 'salt', 'tomatoes', 'ground black pepper', 'thyme', 'eggs', 'green tomatoes', 'yellow corn meal', 'milk', 'vegetable oil']
20130	filipino	['eggs', 'pepper', 'salt', 'mayonaise', 'cooking oil', 'green chilies', 'grilled chicken breasts', 'garlic powder', 'yellow onion', 'soy sauce', 'butter', 'chicken livers']
22213	indian	['water', 'vegetable oil', 'wheat', 'salt']
13162	indian	['black pepper', 'shallots', 'cornflour', 'cayenne pepper', 'onions', 'garlic paste', 'milk', 'butter', 'salt', 'lemon juice', 'water', 'chili powder' , 'passata', 'oil', 'ground cumin', 'boneless chicken skinless thigh', 'garam masala', 'double cream', 'natural yogurt', 'bay leaf']
6602	jamaican	['plain flour', 'sugar', 'butter', 'eggs', 'fresh ginger root', 'salt', 'ground cinnamon', 'milk', 'vanilla extract', 'ground ginger', 'powdered sugar', 'baking powder']
42779	spanish	['olive oil', 'salt', 'medium shrimp', 'pepper', 'garlic', 'chopped cilantro', 'jalapeno chilies', 'flat leaf parsley', 'skirt steak', 'white vinegar', 'sea salt', 'bay leaf', 'chorizo sausage']
3735	italian	['sugar', 'pistachio nuts', 'white almond bark', 'flour', 'vanilla extract', 'olive oil', 'almond extract', 'eggs', 'baking powder', 'dried cranberries']
16903	mexican	['olive oil', 'purple onion', 'fresh pineapple', 'pork', 'poblano peppers', 'corn tortillas', 'cheddar cheese', 'ground black pepper', 'salt', 'iceberg lettuce', 'lime', 'jalapeno chilies', 'chopped cilantro fresh']
12734	italian	['chopped tomatoes', 'fresh basil', 'garlic', 'extra-virgin olive oil', 'kosher salt', 'flat leaf parsley']

```
”id”: 13172,
”cuisine”: ”italian”,
”ingredients”: [
”vodka”,
”marinara sauce”,
”olive oil”,
”crushed red pepper flakes”,
”cream”,
”basil leaves”,
”parmesan cheese”,
”penne pasta”
```

[1] https://www.kaggle.com/competitions/whats-cooking/data

]

Fig 1 shows the number of recipes per cuisine. Again, the dataset skews toward Italian cuisine since it has 7838 recipes. On the other hand, it has low skew toward Brazilian because it has only 467 recipes.

Additionally, cuisines can be distinguished by their unique ingredients, as shown in Fig 2. The deeper analysis illustrates that there are common ingredients used in all cuisines. Fig 3 explains the top 10 common ingredients. Salt, Onion, and Olive Oil are the most common ingredients. Salt is an essential ingredient in almost half of the recipes; 20% contain Onion and Olive Oil.

Each type of cuisine includes a selection of particular components in its dishes. For instance, you can distinguish a Mexican food from others by using avocado, corn tortillas, and salsa. On the other hand, garam masala and ground turmeric most likely represent an Indian dish.

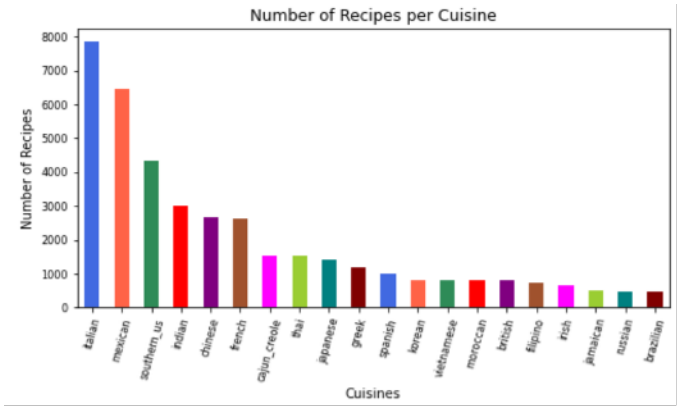

Fig. 1. Cuisine distribution.

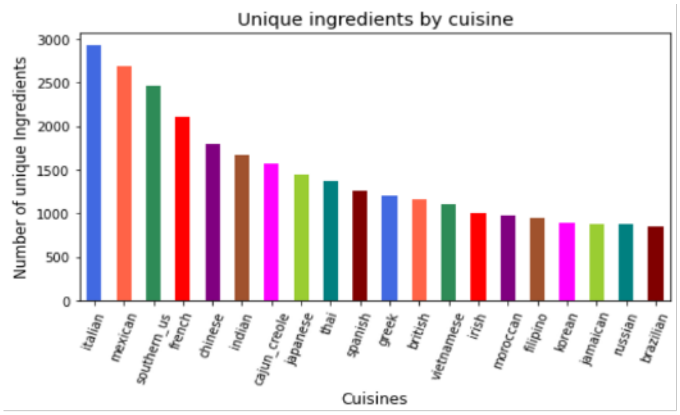

Fig. 2. Unique ingredients by cuisine.

A. Data Preprocessing

Preprocessing is an essential part of exploring the dataset where cleaning processes such as removing the noise data or finding the missing values are compulsory to apply machine learning models. Tokenization is another preprocessing method that divides a phrase, sentence, paragraph, or text document into words or terms. Each of these terms is called a token. Words, numbers, or punctuation could be tokens. In this paper, we applied four different techniques as follows:

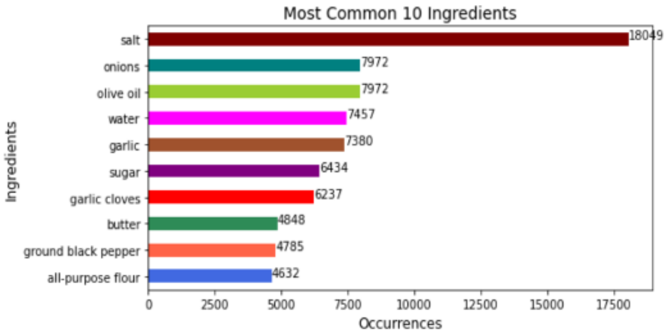

Fig. 3. Top 10 Ingredients.

I. Remove some words from the ingredients. Some ingredients include descriptive words that describe some attributes of the underlying ingredient, such as ”chopped onions” vs. ”onions” ”chopped” is a descriptive word that does not add any value to the learning.

II. Phrases Mapping, another case is considered ”green onion” and ”onion”. The ”green” word in the ingredient name was not removed because it is part of the ingredient name. Therefore, the word ”green onion” mapped to ”onion”.

III. Plural to Singular Mapping, we transformed all of the plurals into their corresponding singular forms. For instance, ”steaks” became ”steak” when it was shortened.

IV. Remove and Replace Space; we decided to remove the unnecessary spaces at the beginning and end of the sentence and replace them with just one space. After completing the cleaning process, we end up with 2929 features.

The features will result in a sparse matrix since each ingredient represents an individual feature [10].A matrix primarily made of zero values is referred to as a sparse matrix... Thus, each ingredient is extracted as a unique attribute. Sparse design matrix (i, j), where i represents the number of recipes and j represents the number of ingredients, shows the sparse representation of each ingredient in the recipes. Moreover, the matrix (i, 1) represents cuisine with values ranging from 0 to 19. A TF-IDF vectorized design matrix approach is utilized to extract features and convert textual data into numerical data [11].

TF-IDF Vectorizer, also known as TfidfVectorizer, is synonymous with CountVectorizer, which is then followed by TfidfTransformer. After transforming a group of unprocessed documents into a matrix of TF-IDF features, it provides floats as output and a score as input [12].

In addition, an unsupervised machine learning technique known as K-means clustering was utilized in order to classify the various cuisines. Since vectorizing the ingredients for each cuisine is an essential process, applied Principal Component Analysis (PCA) is applied to reduce the TF-IDF matrix into two dimensions. PCA is a technique that helps to reduce large dimensions data into small dimensions [13].

Fig 4 shows the cuisines distribution in 3 clusters. You can observe that cuisines are clustered according to the geographical history that they share. The first cluster includes Japanese, Chinese, Korean, Vietnamese, Thai, and Filipino, which are Asian cuisines. The second cluster shows similarities between British, Irish, and Southern US food. Obviously, Italian, Spanish, and Greek cuisines use similar ingredients in the last cluster. Finally, you can notice that the previous cluster includes cuisines from South America, Europe, and the Middle East. Thus, the cuisines are clustered by the ingredient that they share.

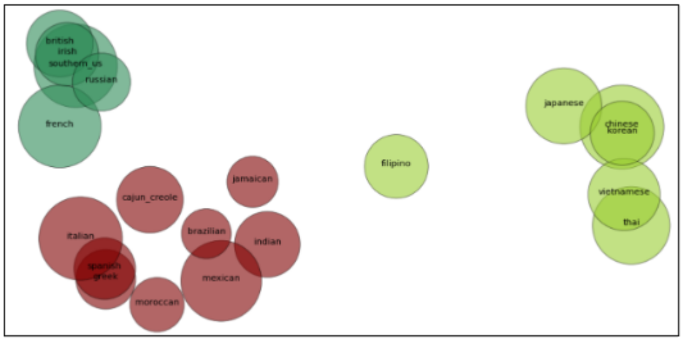

Fig. 4. Visualization of cuisines after PCA and k-Means.

IV. Methodology

In order to categorize the recipes, a multi-class classification system is utilized to predict the two cuisines. Ten different machine learning algorithms are used to develop prediction models. Then those models' performances are compared in terms of accuracy in order to identify the models that are the most accurate. The accuracy of the model's outputs is the metric that is used to evaluate its overall performance.

$$Accuracy = \frac{Number\ of\ correctly\ predicted\ cuisines}{Total\ number\ of\ predictions}$$

These classifiers include K-Nearest Neighbors (KNN), Naïve Bayes (NB), Logistic Regression (LR), Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), Stochastic Gradient Descent (SGD), Gradient Boosting (GB), XGB, and AdaBoost classifiers.

A. KNN: It is an algorithm for pattern recognition that memorizes and acquires knowledge from training data points by computing how each data point correlates to other data points inside an n-dimensional space. In addition, it looks for the k data points in an unknown set of data that are the most closely related to one another [14].

B. NB: An algorithm based on naive Bayes calculates the possibility that any given data point falls into one or more categories (or not) [15]. It is used to calculate the probability of each cuisine for a given recipe and then output the highest probability.

C. LG: With logistic regression, the output is transformed into a probability value by utilizing the logistic sigmoid function. The probability value can then be mapped to two or more variables. [16].

D. RF: It is a classification algorithm that constructs a multitude of decision trees during the training process and outputs the class based on the mode of the individual trees during the training process [17].

E. DT: It builds a model like the structure of the tree. The method uses the if-then rule of mathematics to create subcategories that fit within broader categories and allow for precise, organized categorization [18].

F. SVM: Data is classified, and models are trained using SVM within super finite polarity degrees, producing a 3-dimensional classification [19].

G. SGD: SGD is a simplification. Each iteration calculates En(fwgradient)'s based on a single random example. Each iteration of the stochastic process picks random samples. Despite this shortened procedure's noise, it behaves as expected. Since the stochastic algorithm doesn't need to recall previous samples, it can process them on the fly. Since examples are randomly picked from the ground truth distribution, stochastic gradient descent maximizes anticipated risk [20].

H. GB: The gradient boosting technique is used in regression and classification tasks. Typically, Gradient boosting machines are a class of successful machine-learning algorithms. They can be learned with varied loss functions to suit the application [21].

I. XGBoost: It implements gradient boosted decision trees with an emphasis on speed and performance. XGBoost is an algorithm that has recently dominated the field of applied machine learning and Kaggle competitions in the field of structured or tabular data [22].

J. AdaBoost: Adaptive Boosting is a technique that can be used for both classification and regression problems. It worked by giving a heavier weight to instances that were difficult to classify and less weight to instances that were already handled well. Furthermore, Adaboost combines multiple classifiers (strong and weak) to improve accuracy. Decision stumps, the decision trees with a single split, are the weak learners in AdaBoost [23].

V. Results

The most common two techniques are used to report the accuracy of the models. First, we started the evaluation with ten classifiers using 10-fold cross-validation. Next, the data has been divided into a train-test split technique where 70% of the dataset is used for training while 30% is used for testing with 8888 random states. During the evaluation process, we took a step to optimize the models' hyperparameters by applying the GridSearchCV technique. GridSearchCV is a library function in the "sklearn.model selection package" that helps to fit the estimator (model) on the training set by looping through the given hyperparameters. Moreover, performance was tested through additional cross-validation, testing, and confusion matrix measures.

TABLE II
Models evaluation using 10-Fold cross-validation

Model	Hyperparameters	Accuracy
KNN	leaf_size= 20, metric= 'minkowski', n_neighbors= 10, p= 2, weights= 'distance'	76%
Naïve Bayes	alpha= 0.01	74%
Logistic Regression	C=4.5, max_iter=100, n_jobs=-1, multi_class='ovr',	79%
Random Forest	n_estimators=100, criterion = 'gini', max_features='sqrt'	76%
Decision Tree	Default, criterion='gini', min_samples_split =2	64%
SVM	C= 10, kernel= 'rbf', gamma= 1	81%
SGD	alpha= 0.00001, epsilon= 0.1	74%
Gradient Boosting	Default loss='deviance', learning_rate= 0.1,	74%
XGBoost	learning_rate=0.1, max_depth=5, n_estimators=500, subsample=0.5, colsample_bytree=0.5, eval_metric='auc'	79%
AdaBoost	n_estimators=100, learning_rate= 0.6	60%

Table II shows the performance evaluation of ten classifiers. The K-fold cross-validation method was applied in this experiment to train on multiple train-test splits using 10 cross-validations. The SVM classifier yields achieved the best prediction and the highest accuracy score of 81%, followed by XGBoost, SGD, and Logistic Regression with an accuracy score of 79%. Despite AdaBoost and Decision Tree, all other classifiers obtained accuracy between 74% and 81%. In contrast, Decision Tree achieved 64%, and AdaBoost reached the lowest accuracy, 60%.

Table III demonstrates the achievement of the second experiment. In the second experiment, the train-test split technique was used. As mentioned, the dataset was split into a training set of (70%) and a testing set of (30%). As a result, the SVM algorithm achieved the highest accuracy, with 80%. All other classifiers achieved accuracy between 73% and 78%, except Decision Tree achieved 62%, and AdaBoost achieved the lowest accuracy, 60%. This indicates that with the train-test split, the accuracy has decreased slightly. However, the reduced amount of training data results in reduced calculation costs.

Also, we evaluated the models' performance by calculating the residual Histogram, as shown in Fig 5. It appears that the prediction is close to the actual label. Moreover, we evaluated the performance of the models by generating a confusion matrix, as shown in Fig 6.

From these results, it is clear that SVM performs better prediction in classifying Indian, Mexican, British, and Brazilian food cuisine categories from the dish ingredients. Furthermore, as expected, the SVM provides a much better performance compared to other algorithms.

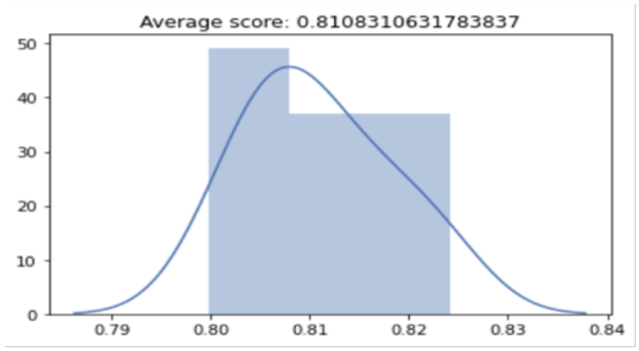

Fig. 5. Evaluation score of SVM.

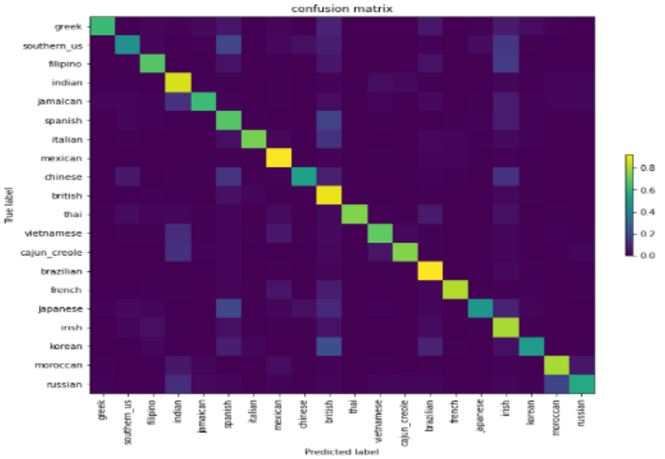

Fig. 6. Confusing matrix of SVM.

VI. Conclusion

This study uses a dataset from Kaggle that includes the ingredients of various dishes worldwide with approximately 40K recipes. The exploration of this dataset through exploratory data analysis (EDA) applying data visualization has been done initially. Then, a preprocessing step was performed to clean and prepare the data to fit machine learning models. Since the dataset contains textual data, a TF-IDF vectorized design matrix approach was used to extract features and convert textual into numerical data before fitting the models. After evaluating ten classifiers using K-fold cross-validation and parameter tuning by applying the

TABLE III
Models performance after train(70%)-test(30%) split

Model	Hyperparameters	Accuracy
KNN	leaf_size= 20, metric= ’minkowski’, n_neighbors= 10, p= 2, weights= ’distance’	75%
Naïve Bayes	alpha= 0.01	73%
Logistic Regression	C=4.5, max_iter=100, n_jobs=-1, multi_class=’ovr’,	78%
Random Forest	n_estimators=100, criterion = ’gini’, max_features=’sqrt’	74%

Decision Tree	Default,
	criterion=’gini’,
	min_samples_split =2
		62%

SVM	C= 10, kernel= ’rbf’, gamma= 1	80%
SGD	alpha= 0.00001, epsilon= 0.1	78%
Gradient Boosting	Default loss=’deviance’, learning_rate= 0.1,	73%
XGBoost	learning_rate=0.1, max_depth=5, n_estimators=500, subsample=0.5, colsample_bytree=0.5, eval_metric=’auc’	78%
AdaBoost	n_estimators=100, learning_rate= 0.6	60%

GridSearchCV technique, the study observed that SVM yields the best prediction for this problem. It can be concluded from the results that SVM, XGBoost, SGD, and Logistic Regression performed well in predicting the cuisines from their recipes when k-fold cross-validation is used. While in the train-test split, SVM is still the best option. The dataset’s wide pattern explains why these classifiers perform better. This study can be further utilized to improve the taste of food items, improve their quality attributes and develop more recipes.

Acknowledgment

I would like to express my appreciation to Jazan University for all of their support.

References

[1] Ueda, M., Takahata, M. & Nakajima, S. User’s food preference extraction for personalized cooking recipe recommendation. *Workshop Of ISWC*. pp. 98-105 (2011)

[2] Ueda, M., Takahata, M. & Nakajima, S. Recipe recommendation method based on user’s food preferences. *Proceedings Of The IADIS International Conference On E-Society*. pp. 591-594 (2011)

[3] Xie, H., Yu, L. & Li, Q. A hybrid semantic item model for recipe search by example. *2010 IEEE International Symposium On Multimedia*. pp. 254-259 (2010)

[4] Forbes, P. & Zhu, M. Content-boosted matrix factorization for recommender systems: experiments with recipe recommendation. *Proceedings Of The Fifth ACM Conference On Recommender Systems*. pp. 261-264 (2011)

[5] Wang, L., Li, Q., Li, N., Dong, G. & Yang, Y. Substructure similarity measurement in chinese recipes. *Proceedings Of The 17th International Conference On World Wide Web*. pp. 979-988 (2008)

[6] Svensson, M., Höök, K. & Cöster, R. Designing and evaluating kalas: A social navigation system for food recipes. *ACM Transactions On Computer-Human Interaction (TOCHI)*. **12**, 374-400 (2005)

[7] Su, H., Lin, T., Li, C., Shan, M. & Chang, J. Automatic recipe cuisine classification by ingredients. *Proceedings Of The 2014 ACM International Joint Conference On Pervasive And Ubiquitous Computing: Adjunct Publication*. pp. 565-570 (2014)

[8] Jayaraman, S., Choudhury, T. & Kumar, P. Analysis of classification models based on cuisine prediction using machine learning. *2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)*. pp. 1485-1490 (2017)

[9] Kalajdziski, S., Radevski, G., Ivanoska, I., Trivodaliev, K. & Stojkoska, B. Cuisine classification using recipe’s ingredients. *2018 41st International Convention On Information And Communication Technology, Electronics And Microelectronics (MIPRO)*. pp. 1074-1079 (2018)

[10] Brownlee, J. A Gentle Introduction to Sparse Matrices for Machine Learning. *Linear Algebra, Mar.* **14** pp. 20 (2018)

[11] Bengfort, B., Bilbro, R. & Ojeda, T. Applied text analysis with python: Enabling language-aware data products with machine learning. (” O’Reilly Media, Inc.”,2018)

[12] Team, A. Amazing Applications/Uses of Data Science Today. *Web. September.* **21** pp. 2015 (13)

[13] Roweis, S. EM algorithms for PCA and SPCA. *Advances In Neural Information Processing Systems*. **10** (1997)

[14] Zhang, M. & Zhou, Z. ML-KNN: A lazy learning approach to multi-label learning. *Pattern Recognition*. **40**, 2038-2048 (2007)

[15] Rish, I. & Others An empirical study of the naive Bayes classifier. *IJCAI 2001 Workshop On Empirical Methods In Artificial Intelligence*. **3**, 41-46 (2001)

[16] Wright, R. Logistic regression.. (American Psychological Association,1995)

[17] Belgiu, M. & Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal Of Photogrammetry And Remote Sensing*. **114** pp. 24-31 (2016)

[18] Song, Y. & Ying, L. Decision tree methods: applications for classification and prediction. *Shanghai Archives Of Psychiatry*. **27**, 130 (2015)

[19] Noble, W. What is a support vector machine?. *Nature Biotechnology*. **24**, 1565-1567 (2006)

[20] Bottou, L. Large-scale machine learning with stochastic gradient descent. *Proceedings Of COMPSTAT’2010*. pp. 177-186 (2010)

[21] Natekin, A. & Knoll, A. Gradient boosting machines, a tutorial. *Frontiers In Neurorobotics*. **7** pp. 21 (2013)

[22] Chen, T. & Guestrin, C. Xgboost: A scalable tree boosting system. *Proceedings Of The 22nd Acm Sigkdd International Conference On Knowledge Discovery And Data Mining*. pp. 785-794 (2016)

[23] Hastie, T., Rosset, S., Zhu, J. & Zou, H. Multi-class adaboost. *Statistics And Its Interface*. **2**, 349-360 (2009)