Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 12, No. 4, December 2024, pp. 940∼949
ISSN: 2089-3272, DOI: 10.52549/ijeei.v12i4.5809

A Diet Recommendation System using TF-IDF and Extra Trees Algorithm

Boudjemaa Boudaa[1,2], Aissa Hammadi[1], Kada Akermi[3]
[1]Computer Science Department, University of Tiaret, Tiaret, Algeria
[2]LabRI-SBA Lab., Ecole Superieure en Informatique, Sidi Bel Abbes, Algeria
[3]Computer Science Department, University of Science and Technology "Mohamed Boudiaf" USTO, Oran, Algeria

Article Info

Article history:

Received Aug 8, 2024
Revised Nov 7, 2024
Accepted Nov 30, 2024

Keywords:

Diet recommendation
Healthy food
Content-based filtering
TF-IDF
Extra Trees algorithm

ABSTRACT

Across the globe, there is a growing emphasis on health and lifestyle choices. However, refraining from unhealthy foods and staying active are not enough; maintaining a well-balanced diet is also essential. Recently, recommendation systems have focused on promoting healthy eating habits, tailoring suggestions for balanced diets based on some parameters like age, gender, height, weight, age, BMI and BMR. Pairing a nutritious diet with regular physical activity can aid in reaching and sustaining a healthy weight, reducing the likelihood of chronic ailments such as heart disease, and enhancing overall well-being. The present paper introduces a novel approach for constructing dietary recommendations with optimized calorie intake, using content-based filtering with the TF-IDF statistical method and machine learning with the Extra Trees algorithm. This approach can generate a dynamic diet based on the calories a person burns and other parameters including the current Body Mass Index (BMI) and BMR (Basal Metabolic Rate). The proposed approach has been tested on a new real-world diet dataset, showcasing its effectiveness in providing diverse and accurate diet recommendations compared to another content-based filtering method and other machine learning algorithms.

Corresponding Author:

Boudjemaa Boudaa
Computer Science Department, University of Tiaret, Tiaret, Algeria
LabRI-SBA Lab., Ecole Superieure en Informatique, Sidi Bel Abbes, Algeria
N°50, Cité 60 logements, rue d'Ain-Bouchekif, Tiaret, Algeria
Email: boudjemaa.boudaa@univ-tiaret.dz

1. INTRODUCTION

Nutrition forms the foundation of human health and development, spanning from the earliest stages of prenatal growth to old age. Healthy eating and proper nutrition are essential for survival, physical development, cognitive growth, performance, productivity, health, and overall well-being.

Nowadays, many human beings suffer from a variety of health issues that might endanger them, many of which are caused by eating unsuitable and harmful meals. The World Health Organisation (WHO) reports that the consumption of unhealthy diets is a primary factor contributing to health problems, particularly non-communicable diseases (NCDs) such as heart disease, stroke, cancer, diabetes, and chronic lung disease. These diseases, collectively responsible for 74% of global deaths, highlight the significance of dietary choices in maintaining overall health. WHO confirms that the proportion of deaths attributed to NCDs has risen to nearly three-quarters of all deaths annually. If this trend persists, it is estimated to escalate to approximately 86%

globally by 2048, equating to 77 million out of 90 million deaths[1].

Several systems and applications have been proposed in the literature to help people eat healthy food and prevent chronic diseases. Recently, recommendation systems which have proved their effectiveness mainly in e-commerce have specialised in food and diet. In this field, we distinguish between food recommender system (FRS) and diet recommender system (DRS) which both aim to enhance overall well-being, but with different objectives, scope, and the type of recommendations they offer. An FRS focuses on enhancing the user's dining experience and culinary exploration by suggesting food items to users based on their preferences, past behaviour, or other relevant factors. Despite this, a DRS prioritise supporting users in achieving optimal nutrition, health outcomes, and dietary adherence by recommending personalised dietary plans or nutritional strategies to users based on their health goals, dietary preferences, lifestyle, and other relevant factors.

While the literature witnessed an abundance in terms of works on FRS [1], there are only just a few on DRS. Here, we can mention the following relevant works. To help manage the diets of hypertensive individuals, the work in [2] proposed a DASH diet recommendation system that utilizes machine learning and content-based filtering techniques to suggest healthy menus and dishes. In more depth, the authors in [3] introduced a deep learning approach based on Long Short-Term Memory (LSTM) that incorporates health-related medical data collected through an Internet of Medical Things (IoMT) framework. This approach automatically allows for the identification of suitable diets for patients, considering their disease and other characteristics like age, gender, weight, calorie intake, protein, fat, sodium, fibre, and cholesterol levels. Also based on machine learning (ML), a diet recommendation system is provided in [4]. It takes into account both current and past food consumption records, providing more reliable diet recommendations. In [5], the authors introduced a framework that combines machine learning techniques for big data analytics with natural language processing (NLP). This framework employs an algorithm called Intelligent Recommender for Healthy Diet (IR-HD) to analyze data and offer healthy diet recommendations.

Unlike previous diet models that focused solely on nutritional information or addressed a specific issue faced by the user, the authors in [6] proposed a more comprehensive approach. Their model not only covers most nutritional features of food but also emphasizes the fat, carbohydrate, calorie, and protein content. Additionally, it considers the user's Body Mass Index (BMI) when preparing personalized diet plans. In this approach, K-means clustering is employed to group foods with similar nutritional content, and a random forest classifier is then used to develop the model for recommending a diet to the user. In [7], the proposed diet recommendation system employs the K-means clustering algorithm, the food inference algorithm, and the patient nutrition calculation algorithm to customize the most suitable diet for each individual, taking into account their gender, age, and current nutritional status.

The present paper aims to shed more light on the development of diet recommender systems and propose a DRS to tailor suggestions for balanced diets based on some factors such as age, gender, height, weight, age, BMI and BMR (Basal Metabolic Rate). It seeks to introduce a novel approach for constructing dietary recommendations through content-based filtering via TF-IDF and employing a specific machine-learning algorithm called Extra Trees. This approach has been tested on a new real-world diet dataset, demonstrating its efficacy compared with another content-based filtering technique and other standard machine-learning algorithms.

The remainder of this paper is organized as follows. Section 2 details the development method of our proposed diet recommendation system using the TF-IDF technique and the Extra Trees algorithm. Section 3 discusses the results obtained from extensive experiments, highlighting key insights and future directions. Finally, the paper concludes in Section 4.

2. METHOD

This section presents the development model of our personalised diet recommendation system, which leverages content-based filtering and advanced machine-learning algorithms to offer tailored dietary advises.

2.1. TF-IDF

Recommendation systems (RSs) are pivotal in today's digital landscape, offering suggestions to users across various platforms. Many RSs types exist [8], in which the four primary ones are: Content-based filtering (CBF), Collaborative filtering (CF), Knowledge-based (KB) and Hybrid approaches. Each approach employs

[1]`https://www.who.int/publications/i/item/9789240074323`

distinct techniques to deliver tailored recommendations to users based on their preferences and interactions. In the literature of diet recommendations [9, 10], content-based filtering has not been well exploited. This has inspired the current work to leverage the CBF's capabilities for recommending accurate diets.

Content-based filtering is a recommendation system method that suggests items to users based on the features of items they have liked in the past. It processes by comparing extracted features in the user profile on items and current item's profile. The user profile is constructed on the user's item preferences by aggregating the features of those items. Information about the user's item preferences (e.g., some diets) can be gathered explicitly (i.e., rating) or implicitly (i.e., actions). Here, only the user's information is required, rather than information from other similar users (as in collaborative filtering), so only a small scope of information is needed for recommendations by CBF [11]. Furthermore, CBF avoids the cold start problem, and its recommendations are highly relevant and transparent to the user. Since most of dietary data work with text data, we will utilise the TF-IDF (Term Frequency-Inverse Document Frequency) method to introduce our recommendation model into a content-based filtering system.

The TF-IDF method is used to evaluate the importance of words in a textual corpus. The importance is proportional to the number of times the words appear in the document and inversely proportional to the frequency of words appearing in the corpus.

- TF represents the frequency of words, indicating the number of times they appear in a corpus (Eq. 1). This consists of calculating the number of word appearances out of the total number of words present in the corpus [12].

$$TF = \frac{Term\ Frequency}{Total\ number\ of\ terms\ in\ document} \tag{1}$$

- IDF quantifies the significance of a term across the entire corpus by computing the logarithm of the inverse of the proportion of documents containing the term (Eq. 2). This involves dividing the total number of documents in the corpus by the number of documents where the term appears and taking the logarithm of the result [12].

$$IDF = \log\left(\frac{Total\ number\ of\ documents\ in\ corpus}{Number\ of\ documents\ containing\ the\ term}\right) \tag{2}$$

Recently, advanced techniques have been developed for creating content-based recommendation systems, incorporating sophisticated machine learning algorithms. These algorithms typically require numerical data, so when working with textual information, it is necessary to convert the text into numerical format with a process called vectorization. One common approach is TF-IDF (Term Frequency-Inverse Document Frequency) vectorization. This method calculates the TF-IDF score for each word in a document relative to the entire corpus. Each document is then represented by a vector, where each element corresponds to the TF-IDF score of a particular word across the entire set of documents. With these vectors, you can perform various tasks, such as determining the similarity between two documents by comparing their TF-IDF vectors using methods like cosine similarity. This approach is also useful for various machine learning classifiers.

2.2. Extra Trees Algorithm

Machine learning (ML) algorithms are computational techniques that enable systems to learn from data, identify patterns, and make informed decisions with minimal human intervention. They can be broadly classified into several categories: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Each category serves different purposes and is suited to specific types of tasks. ML algorithms find applications across various domains, including healthcare (such as predicting disease outbreaks), finance (such as detecting fraud), marketing (such as segmenting customers), technology (such as speech recognition and image classification), and recommendation systems (such as offering personalized suggestions).

In recommendation systems, ML algorithms help provide tailored and relevant suggestions to users by analysing their preferences, behaviours, and past interactions. In this study, we have chosen to use the Extra Trees algorithm, an ensemble supervised learning method based on decision trees. The Extra Trees algorithm, short for Extremely Randomised Trees, is similar to the random forests algorithm. It involves constructing multiple decision trees and combining their outputs to improve prediction accuracy. For each tree, the algorithm randomly samples data without replacement, creating a unique dataset. It also selects a random subset of features for each tree. A key characteristic of the Extra Trees algorithm is its use of random splits: instead of determining the optimal split point based on criteria like Gini impurity or entropy, the algorithm randomly selects a splitting value. This randomness in splitting increases diversity among the trees, leading to less correlation and more robust predictions [13].

The Extra Trees algorithm aggregates the predictions from its multiple decision trees by averaging their outputs. This ensemble approach enhances the overall accuracy and robustness of the model by mitigating over-fitting and effectively capturing complex patterns within the data.

2.3. A Proposed Model for DRSs

We follow a structured approach involving several key components to design the proposed diet recommendation system. As depicted in Figure 1, we initially gather a comprehensive dataset comprising dietary information and user preferences. This raw data undergoes preprocessing actions by deleting the null and duplicate values. We then apply a content-based approach using TF-IDF method (see 2.1.) to analyze the attributes of dietary items and identify patterns. This is followed by the application of Extra Trees Machine learning algorithm (see 2.2.) to further classify and refine our recommendations. Finally, the system generates personalised diet plans tailored to individual user needs and preferences, providing practical and health-conscious dietary suggestions.

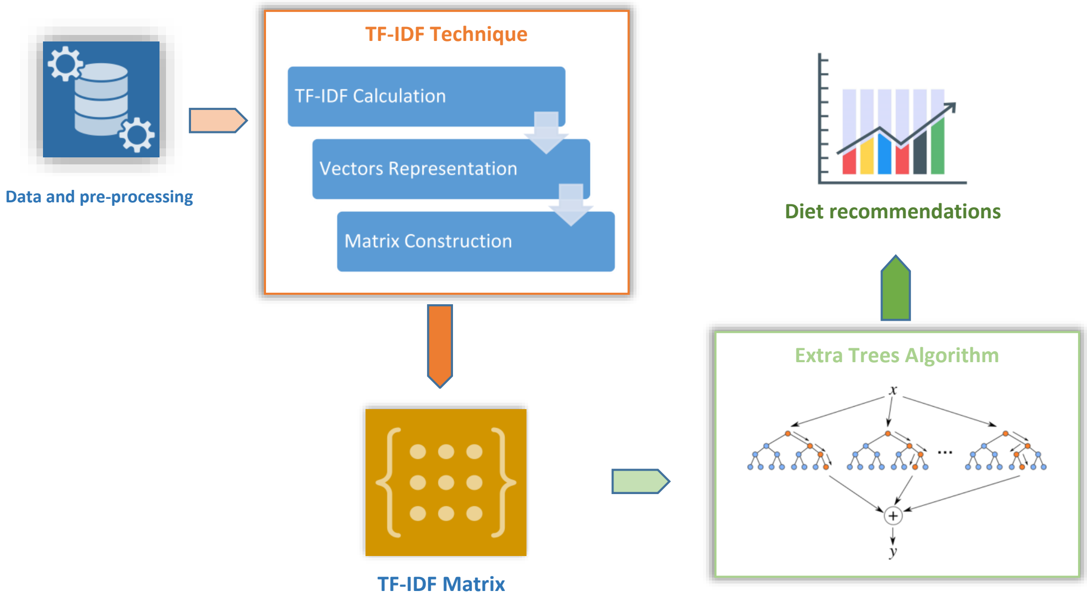


Figure 1. Overview of the Proposed Model Architecture.

3. RESULTS AND DISCUSSION

The obtained results from the experimentation of the proposed model (see 2.3.) will be explained in this section with a comprehensive discussion. Before this, we should exhibit some technical settings that have been used in this experimentation.

3.1. Dataset and preprocessing

The dataset used for this study will be described with the pre-processing actions applied to it.

3.1.1. Dataset

The proposed diet recommendation model harnesses the power of a comprehensive dataset sourced from Kaggle.com[2], boasting an expansive collection of over 80,000 meticulously curated cases. This dataset serves as a treasure trove of insights, meticulously analyzing and cataloging recommended diets tailored to the nuanced needs of diverse body types and individuals. Within this rich repository of information, an array of essential variables is meticulously documented, including age, gender, height, weight, body mass index (BMI), fitness level, fitness goals, medical history, basal metabolic rate (BMR), and caloric intake (see Figure 2). Each of these variables plays a critical role in the DRS, contributing to the holistic understanding of an individual's dietary requirements. At the core of our analysis lies the target column: the recommended diet. This culmination of data and insights allows our system to navigate the intricate landscape of dietary recommendations with precision and finesse.

	Age	Gender	Height	Weight	BMI	Fitness Level	Fitness Goal	Medical History	Diet Recommended	BMR	Calories
0	59	F	149	42	18.928228	Underweight	weight gain	none	High calorie	1051.38	1997.622
1	58	F	149	42	18.928228	Underweight	weight gain	none	High calorie	1056.08	2006.552
2	59	F	152	42	18.178670	Underweight	weight gain	none	High calorie	1057.02	2008.338
3	57	F	149	42	18.928228	Underweight	weight gain	none	High calorie	1060.78	2015.482
4	59	F	149	43	19.378900	Normal weight	muscle building	none	High Protein Diet	1060.84	1644.302

Figure 2. Sample from the dataset.

3.1.2. Data preprocessing

The dataset mentioned in above presents a multifaceted landscape of categorical data, blending numerical values with textual descriptors, thereby offering a nuanced portrayal of dietary recommendations across a diverse spectrum. However, before delving into the realm of model construction, a meticulous preprocessing phase is imperative to ensure data integrity and model efficacy.

Central to this preprocessing endeavor is the identification and handling of null values. If left unaddressed, null values can compromise the robustness of our analyses and introduce biases into our model. Employing a strategic approach, we meticulously scrutinize each feature, implementing techniques such as imputation or removal to mitigate the impact of missing data.

Furthermore, the presence of duplicated values warrants careful attention. Duplicates not only skew statistical analyses but also distort the learning process of machine learning algorithms, leading to suboptimal model performance. Employing deduplication strategies, we systematically identify and eliminate redundant entries, streamlining the dataset for enhanced model building.

Additionally, categorical data, with its diverse array of textual values, demands specialized treatment. Techniques such as one-hot encoding or label encoding are employed to transform textual descriptors into a format conducive to computational analysis, facilitating seamless integration into machine learning pipelines.

Through meticulous preprocessing, we lay the groundwork for the subsequent stages of model development, ensuring that our analyses are founded in a robust foundation of clean and standardized data. By adhering to best practices in data prepossessing, we fortify our model against potential pitfalls and pave the way for reliable insights into dietary recommendations tailored to individual needs.

3.2. Evaluation metrics

To assess the performance of the proposed DRS model, we utilized various evaluation metrics, including:

- **Accuracy:** measures how close a group of predictions is to its actual value [14]. It is calculated as follows (Eq. 3).

$$Accuracy = \frac{Correct\ predections}{All\ predections} \tag{3}$$

- **Precision:** is the proportion of relevant items correctly recommended among all the recommended items [15]. It is represented by the following equation (Eq. 4).

$$Precision = \frac{True\ positives}{True\ positives + False\ positives} \tag{4}$$

[2]https://www.kaggle.com/datasets/salmansiddiquecs/diet-exercise-calorise-recommandation-dataset

- **Recall:** is a measure of completeness that determines the proportion of relevant items recommended out of the total number of relevant items [15]. It is calculated as follows (Eq. 5):

$$Recall = \frac{True\ positives}{True\ positives + False\ negatives} \tag{5}$$

- **F1-score:** is the precision and recall weighted mean [15]. So, the F1-score is a statistic that combines precision and recall which is calculated using the equation below (Eq. 6).

$$F1 - score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

3.3. Implementation

The diet recommendation system leverages a combination of a content-based approach by TF-IDF and machine learning by Extra trees to generate personalized diet suggestions. The process begins with data preparation, in which 'Gender', 'Fitness Level', 'Fitness Goal', and 'Medical History', are consolidated into a unified single text feature (i.e., 'info' column). This consolidation serves as a foundation for the TF-IDF transformation, in which each text is transformed into numerical value using the TF-IDF vectorizer. TF-IDF measures the importance of words in the text based on their frequency/rarity in the document and across the dataset providing a nuanced understanding of their relevance. This transformation allows the model to represent the textual features as a matrix, where each row corresponds to a user profile and each column represents a unique word. The values in this matrix (see Figure 3) reflect the relevance of each word in describing the user's attributes.

In addition to the text-based features, the system also incorporates numerical features such as 'Age', 'Height', 'Weight', 'BMI', 'BMR', and daily calorie intake. These numerical features are combined with the TF-IDF features to create a comprehensive feature matrix (as in Figure 3). The target variable, representing the recommended diet for each user, is then aligned with this feature matrix.

	building	celiac	diabetes	gain	gout	hypercholesterolaemia	loss	muscle	none	normal	obesity	overweight	underweight	weight
0	0.000000	0.000000	0.000000	0.632718	0.000000	0.000000	0.000000	0.000000	0.342385	0.000000	0.000000	0.0	0.632718	0.286545
1	0.000000	0.000000	0.000000	0.632718	0.000000	0.000000	0.000000	0.000000	0.342385	0.000000	0.000000	0.0	0.632718	0.286545
2	0.000000	0.000000	0.000000	0.632718	0.000000	0.000000	0.000000	0.000000	0.342385	0.000000	0.000000	0.0	0.632718	0.286545
3	0.000000	0.000000	0.000000	0.632718	0.000000	0.000000	0.000000	0.000000	0.342385	0.000000	0.000000	0.0	0.632718	0.286545
4	0.530633	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.530633	0.302196	0.530633	0.000000	0.0	0.000000	0.252911
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
84575	0.000000	0.759883	0.000000	0.000000	0.000000	0.000000	0.369181	0.000000	0.000000	0.000000	0.502215	0.0	0.000000	0.184559
84576	0.000000	0.000000	0.759883	0.000000	0.000000	0.000000	0.369181	0.000000	0.000000	0.000000	0.502215	0.0	0.000000	0.184559
84577	0.000000	0.000000	0.000000	0.000000	0.759883	0.000000	0.369181	0.000000	0.000000	0.000000	0.502215	0.0	0.000000	0.184559
84578	0.000000	0.000000	0.000000	0.000000	0.000000	0.759883	0.369181	0.000000	0.000000	0.000000	0.502215	0.0	0.000000	0.184559
84579	0.000000	0.759883	0.000000	0.000000	0.000000	0.000000	0.369181	0.000000	0.000000	0.000000	0.502215	0.0	0.000000	0.184559

84580 rows × 14 columns

Figure 3. An Excerpt from the Resulting TF-IDF Matrix.

The data is split into training and testing sets, with the training set used to train the Extra Trees machine learning classifier. Trained on this dataset, our model gains the capability to discern patterns and make informed predictions by both textual semantics and numerical trends. Figure 4 shows an excerpt on how this algorithm builds its classification to predict accurate diet recommendations.

Finally, the trained model is evaluated on the testing set using accuracy, precision, recall, and F1-score metrics. These metrics assess the model's performance in correctly recommending diets based on the input features.

By combining the power of TF-IDF for textual data representation and the Extra Trees classifier for prediction, the system can effectively provide personalised diet recommendations to users.

3.4. Results and comparison

To prove the effectiveness of the proposed diet recommendation model, we have conducted a comparison in two different scenarios on the two main components in its architecture. In the first scenario, the performance of the Extra trees algorithm has been compared with other popular state-of-the-art machine learning algorithms such as SVM (Support Vector Machine), KNN (k-nearest neighbours), DT (Decision tees) and RF (Random forest). However, in the second, TD-IDF has been replaced by Cosine Similarity, another technique widely used in content-based recommendation systems. After experimenting with the proposed model,

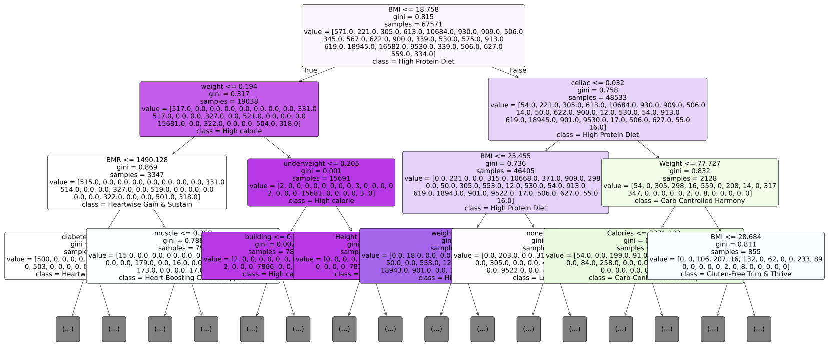


Figure 4. Sample of Extra Trees Algorithm Calculation on TF-IDF Matrix.

Table 1 shows obtained results for the first scenario. However, in Table 2, we find the results of the second one. Figures 5 and 6 illustrate these results respectively. These outcomes have been evaluated based on the previously mentionned metrics, including accuracy, precision, recall, and F1-score.

Table 1. Performance of TF-IDF with Machine Learning Algorithms.

Algorithm	Accuracy	Precision	Recall	F1-Score
SVM	0.8206	0.6738	0.8206	0.7399
KNN	0.7865	0.6766	0.7865	0.7273
Decision Tree	0.9853	0.9854	0.9853	0.9853
Random Forest	0.9882	0.9884	0.9882	0.9882
Extra Trees	**0.9915**	**0.9916**	**0.9915**	**0.9915**

Table 2. Performance of Cosine Similarity with Machine Learning Algorithms.

Algorithm	Accuracy	Precision	Recall	F1-Score
SVM	0.8216	0.6753	0.8216	0.7412
KNN	0.7836	0.7785	0.7836	0.7806
Decision Tree	0.7740	0.8070	0.7740	0.7900
Random Forest	0.7813	0.7761	0.7813	0.7787
Extra Trees	0.7831	0.8067	0.7831	0.7946

3.5. Discussion

Through rigorous experimentation and analysis, we have gained insights into the model's performance across different scenarios. In terms of the used evaluation metrics (accuracy, precision, recall, F1-score), we observed in Figure 5 that our diet recommendation model with Extra trees algorithm takes a clear superiority against SVM and KNN algorithms. However, it is slightly high against DT and RF. This is due that Extra trees and RF are wholly based on using decision trees (DT). As in Random forest, the Extra tree algorithm combines the output of multiple decision trees to reach a single result.

In contrast, Figure 6 confirms that the Cosine Similarity technique shows low results against TF-IDF for the proposed diet recommendation.

The TF-IDF technique serves as a pivotal component in this model, enabling it to extract meaningful signals from textual data while maintaining the context and significance of each term. In addition, the performance of the Extra Trees algorithm in machine learning enables the delivery of accurate diet recommendations by leveraging the combined insights from both textual and numerical data.

The proposed diet recommendation system enhances the effectiveness of combining the content-based approach and machine learning to generate personalised diet suggestions for health and well-being..

3.6. Limitations and future directions

Despite the strengths discussed above, it is important to highlight some limitations of the proposed model:

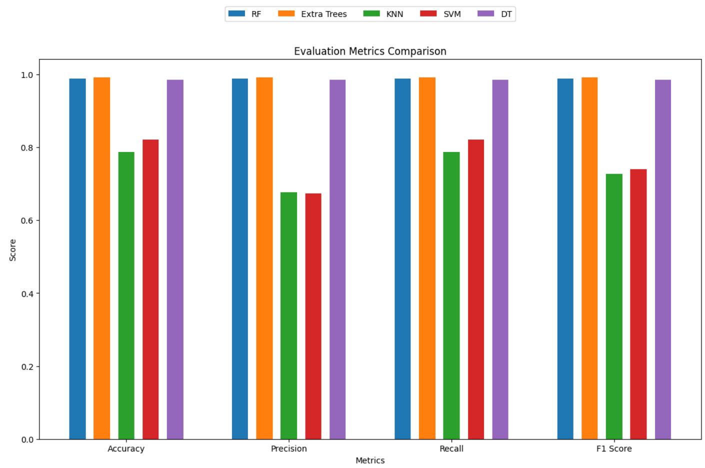


Figure 5. Model's performance using different machine learning algorithms and TF-IDF.

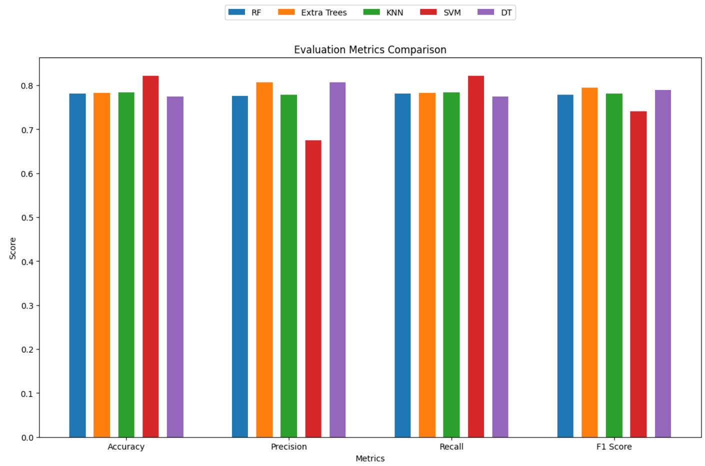


Figure 6. Model's performance using different machine learning algorithms and cosine similarity.

- **Diet-Only Focus**: The current model exclusively provides diet plans without specific food recommendations, limiting its ability to suggest detailed meal options that align with users' dietary needs and preferences.
- **Lack of Long-Term Tracking**: The model lacks the capability to adjust recommendations over time, restricting its capability to adapt to the evolving needs of users.
- **Computational Complexity**: Advanced algorithms like Extra Trees can be resource-intensive, causing delays, especially on low-performance devices or with large datasets.

This study on diet recommendation systems also identifies issues that could open avenues for future research directions:

- **Data Quality Dependence**: The DRS's accuracy relies heavily on the quality of the input data. Incomplete or biased datasets can lead to inaccurate recommendations, especially for specific groups like elderly or chronic disease patients. This provokes the need to handle user input variability or inaccuracies in self-reported dietary habits.
- **Cultural and Dietary Preference Expansion**: The DRS models should handle extended datasets that include diverse cultural diets and preferences for more personalized recommendations.
- **Clinical Validation**: There is a necessity to conduct clinical studies to validate and refine the model's dietary recommendations based on real-world outcomes.

- **Dynamic Learning and Dataset Enhancement**: An auto-learning mechanism is required to enable the model to adapt to new data over time. Additionally, the diets dataset should be expanded to include a broader range of cases, covering diverse dietary needs and specific food recommendations for enhanced personalization.

4. CONCLUSION

This paper has outlined a model for building a personalised dietary recommendation system by utilizing TF-IDF and Extra trees algorithm. We have demonstrated the importance of this model by way of extensive experimentation and comparison with main machine learning classifiers. Through our analysis, we have highlighted the strengths of a content-based approach combined with machine learning algorithms for delivering accurate and personalised dietary advice. The integration of numerical and textual data has significantly enhanced our model's capabilities, resulting in more precise recommendations.

Our findings underscore the potential of machine learning in analyzing patterns in dietary guidance and providing more reliable, accurate, personalized, and context-aware diet recommendations than the simple calculations offered by medical and health applications based solely on a few parameters. Machine learning, and even deep learning, can pave the way for more individualized and dependable health solutions.

For a future work, our research will focus on two areas. First, we will concentrate on employing deep learning techniques for diet recommendations, particularly when managing large dietary datasets. Second, we plan to implement some of the key works from the literature review for eventual comparison with our proposed model.

REFERENCES

[1] J. N. Bondevik, K. E. Bennin, Önder Babur, and C. Ersch, "A systematic review on food recommender systems," *Expert Systems with Applications*, vol. 238, p. 122166, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417423026684

[2] R. Sookrah, J. D. Dhowtal, and S. D. Nagowah, "A dash diet recommendation system for hypertensive patients using machine learning," in *2019 7th international conference on information and communication technology (ICoICT)*. IEEE, 2019, pp. 1–6.

[3] C. Iwendi, S. Khan, J. H. Anajemba, A. K. Bashir, and F. Noor, "Realizing an efficient iomt-assisted patient diet recommendation system through machine learning model," *IEEE Access*, vol. 8, pp. 28 462–28 474, 2020.

[4] M. Geetha, C. Saravanakumar, K. Ravikumar, and V. Muthulakshmi, "Human body analysis and diet recommendation system using machine learning techniques." EAI, 1 2021.

[5] M. A. Lambay and S. P. Mohideen, "A hybrid approach based diet recommendation system using ml and big data analytics," *Journal of Mobile Multimedia*, vol. 18, no. 06, p. 1541–1560, Jul. 2022. [Online]. Available: https://journals.riverpublishers.com/index.php/JMM/article/view/12037

[6] N. Vignesh, S. Bhuvaneswari, K. Kotecha, and V. Subramaniyaswamy, "Hybrid diet recommender system using machine learning technique," in *Hybrid Intelligent Systems*, A. Abraham, T.-P. Hong, K. Kotecha, K. Ma, P. Manghirmalani Mishra, and N. Gandhi, Eds. Cham: Springer Nature Switzerland, 2023, pp. 106–115.

[7] M. Ahmad, A. U. Khan, and M. Sajid, "A diet recommendation system for persons with special dietary requirements," *Journal of Computing amp; Biomedical Informatics*, vol. 5, no. 01, p. 153–164, Jun. 2023. [Online]. Available: https://www.jcbi.org/index.php/Main/article/view/180

[8] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowledge-Based Systems*, vol. 46, pp. 109–132, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705113001044

[9] I. Orue-Saiz, M. Kazarez, and A. Mendez-Zorrilla, "Systematic review of nutritional recommendation systems," *Applied Sciences*, vol. 11, no. 24, 2021. [Online]. Available: https://www.mdpi.com/2076-3417/11/24/12069

[10] M. Shah, S. Degadwala, and D. Vyas, "Diet recommendation system based on different machine learners: A review," in *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*. IEEE, 2022, pp. 290–295.

[11] F. Isinkaye, Y. Folajimi, and B. Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egyptian Informatics Journal*, vol. 16, no. 3, pp. 261–273, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1110866515000341

[12] M. Chiny, M. Chihab, O. Bencharef, and Y. Chihab, "Netflix recommendation system based on tf-idf and cosine similarity algorithms," *no. Bml*, pp. 15–20, 2022.

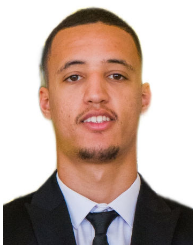
[13] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, Apr. 2006.

[14] R. Potukuchi and P. Upadhyay, *Prediction Model for Precision Agriculture Using Machine Learning*, 01 2024, pp. 627–644.

[15] M. Ahmad, A. U. Khan, and M. Sajid, "A diet recommendation system for persons with special dietary requirements," *Journal of Computing & Biomedical Informatics*, vol. 5, no. 01, 2023.

BIOGRAPHY OF AUTHORS

Boudjemaa Boudaa is an Associate Professor in the Department of Computer Science at Ibn Khaldoun University of Tiaret, Algeria. He received his PhD in Computer Science from Tlemcen University in 2017. He also holds a Magister degree in Computer Science from Tiaret University (2009) and an Engineering degree in Computer Science from Oran University (1995). Currently, he is a member of the "Service-Oriented Computing" research team at the LabRI-SBA Laboratory, ESI Sidi-Bel-Abbes. His research interests include Context-Aware Systems, Recommender Systems, Software Engineering, Autonomous Vehicles, Computer Vision, and Machine Learning. He can be contacted at email: boudjemaa.boudaa@univ-tiaret.dz.



Aissa Hammadi is currently a graduate of the Department of Computer Science at Ibn Khaldoun University of Tiaret, where he received his Bachelor's Degree in Information Systems and Software Engineering in 2024. He plans to continue his studies by pursuing a master's degree in software engineering. Alongside his academic pursuits, he engages in research in the field of artificial intelligence, with a particular interest in recommendation systems and machine learning. He can be contacted at email: aissa.hammadi@univ-tiaret.dz.

Kada Akermi earned his Ph.D. in Computer Science from the University of Science and Technology of Oran (USTO Oran) in 2020. His doctoral research focused on artificial intelligence, autonomous vehicles and control systems. Prior to this, he completed his Master's degree in Computer Science at the University of Ibn Khaldoun Tiaret in 2012. His academic work has significantly contributed to advancements in these fields. Dr. Akermi's expertise encompasses a range of technologies and methodologies related to his research interests. He is available for contact via email at kada.akermi@univ-usto.dz.