IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

Food Quality Identification using Machine Learning

Bushra Ayesha[1], Dr. Narendra G. Bawane[2], Urvashi Agrawal[3]
Student, Department of ETC[1]
Professor, Department of ETC[2,3]
Jhulelal Institute of Technology, Nagpur, Maharashtra, India

Abstract: *In food safety and regulation, there is a need for an automated system to be able to make predictions on which adulterants (unauthorized substances in food) are likely to appear in which food products. An illegal red dye, to adulter "strawberry ice cream", but not "bread". In this work, we show a novel application of deep neural networks in solving this task. We leverage data sources of commercial food products, hierarchical properties of substances, and documented cases of adulterations to characterize ingredients and adulterants. Taking inspiration from natural language processing, we show the use of recurrent neural networks to generate vector representations of ingredients from Wikipedia text and make predictions. Finally, we use these representations to develop a sequential method that has the capability to improve prediction accuracy as new observations are introduced. The results outline a promising direction in the use of machine learning techniques to aid in the detection of adulterants in food.*

Keywords: Adulterants, Deep Neural Networks, Machine Learning.

I. INTRODUCTION

Ensuring the safety and integrity of the food supply chain is paramount for public health and consumer confidence. However, the prevalence of food adulteration, where unauthorized substances are introduced into food products, poses significant challenges to food safety and regulation. Detecting and preventing such adulteration requires timely and accurate identification of potential adulterants and their likely associations with specific food products.

In this work, we address the need for an automated system capable of predicting which adulterants are likely to appear in which food products. We propose a novel application of deep neural networks to tackle this task, leveraging diverse data sources, hierarchical properties of substances, and documented cases of adulterations.

Inspired by techniques from natural language processing, we utilize recurrent neural networks (RNNs) to generate vector representations of ingredients extracted from Wikipedia text. These representations capture semantic relationships and contextual information, enabling us to characterize ingredients and adulterants in a meaningful way.

Furthermore, we introduce a sequential method that leverages the generated vector representations to improve prediction accuracy over time. By incorporating new observations into the model, this method adapts and refines its predictions, enhancing its capability to detect potential adulteration instances.

Our approach represents a significant advancement in the field of food safety and regulation, offering a data-driven and automated solution to the complex problem of adulterant prediction. The results of our study demonstrate the potential of deep neural networks, particularly recurrent architectures, in aiding the detection of adulterants in food products.

Through this research, we aim to contribute to the development of more robust and effective tools for safeguarding the integrity of the food supply chain, ultimately enhancing consumer safety and trust in the food products they consume.

he food industry plays a pivotal role in global health, economy, and social well-being by ensuring the production and distribution of safe, nutritious, and high-quality food products. However, maintaining consistent food quality throughout the supply chain poses significant challenges due to factors such as variations in raw materials, processing methods, and environmental conditions. Ensuring food quality is not only essential for consumer satisfaction but also critical for public health and safety, as substandard or contaminated food can lead to adverse health effects and economic losses.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

Traditional methods of food quality assessment often rely on manual inspection, sensory evaluation, and chemical analysis, which can be time-consuming, subjective, and prone to errors. Moreover, these methods may not be scalable or adaptable to the increasingly complex and dynamic nature of the food industry. As a result, there is a growing need for innovative technologies and approaches to enhance food quality identification processes and mitigate associated risks.

In recent years, machine learning has emerged as a powerful tool for analyzing large datasets, extracting meaningful patterns, and making accurate predictions across various domains. By leveraging advanced algorithms and computational techniques, machine learning offers the potential to revolutionize food quality identification by automating tasks, improving accuracy, and enabling real-time monitoring of food products throughout the supply chain.

In this study, we aim to explore the application of machine learning in food quality identification, with a focus on developing predictive models capable of assessing the quality of food products based on their inherent characteristics and properties. Specifically, we seek to address the following objectives:

- Investigate the feasibility of using machine learning algorithms to analyze diverse datasets related to food composition, sensory attributes, and manufacturing processes.
- Develop and evaluate machine learning models for predicting the quality of food products based on relevant features and attributes.
- Assess the performance of the developed models in terms of accuracy, reliability, and scalability.
- Explore potential applications of machine learning in enhancing food quality assurance practices, including quality control, risk management, and regulatory compliance.

II. LITERATURE SURVEY

- **"Deep Learning Approaches for Food Adulteration Detection":** This study explores the application of deep learning techniques, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), for detecting food adulteration. It discusses various data sources and preprocessing techniques used for training the models and evaluates their performance in detecting known and unknown adulterants.
- **"Automated Food Safety Assurance Using Machine Learning Algorithms":** This paper provides an overview of machine learning algorithms applied to various aspects of food safety, including adulteration detection. It reviews different approaches, such as image processing, spectroscopy, and sensor fusion, and discusses their strengths and limitations in ensuring food quality and integrity.
- **"Predictive Modeling of Food Adulteration Using Semantic Analysis":** This research investigates the use of semantic analysis techniques to predict potential instances of food adulteration. It discusses the extraction of semantic features from textual data sources, such as ingredient lists and product descriptions, and evaluates the effectiveness of different machine learning models in predicting adulteration events.
- **"Advances in Data Mining Techniques for Food Safety and Quality Assurance":** This review article provides an overview of recent advances in data mining techniques applied to food safety and quality assurance. It discusses the use of data mining algorithms for analyzing large scale datasets to identify patterns and trends related to food adulteration and contamination.
- **"Integration of Knowledge Graphs in Food Safety Analysis":** This study explores the integration of knowledge graphs into food safety analysis for identifying potential adulteration patterns and relationships between ingredients and adulterants. It discusses the construction of knowledge graphs from diverse data sources and their use in supporting decision-making processes for food safety management.
- **"Deep Learning for Food Product Authentication":** This research focuses on the application of deep learning methods for authenticating food products and detecting counterfeit or adulterated items. It discusses the use of convolutional neural networks (CNNs) for analyzing product images and spectroscopic data to identify anomalies indicative of adulteration or contamination.
- **"Predictive Modeling of Food Adulteration Using Bayesian Networks":** This study investigates the use of Bayesian networks for building predictive models of food adulteration. It discusses the representation of probabilistic relationships between ingredients and adulterants and evaluates the performance of Bayesian

DOI: 10.48175/IJARSCT-18686

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

network models in predicting potential instances of adulteration. These studies collectively highlight the diverse approaches and methodologies employed in the field of food adulteration detection, ranging from traditional machine learning algorithms to advanced deep learning techniques. They provide valuable insights into the current state of research and identify opportunities for further advancements in ensuring food safety and integrity.

- The study of food quality identification using machine learning techniques builds upon a rich body of research spanning multiple disciplines, including food science, computer science, and data analytics. In this literature survey, we review key studies and methodologies relevant to our research objectives, focusing on three main areas: traditional methods of food quality assessment, applications of machine learning in the food industry, and recent advancements in predictive modeling for food quality identification.
- Traditional methods of food quality assessment have long been established in the food industry and encompass a variety of techniques for evaluating the physical, chemical, and sensory properties of food products. These methods often involve manual inspection, laboratory analysis, and sensory evaluation by trained experts. For example, chemical analysis techniques such as chromatography and spectrophotometry are commonly used to measure the concentration of specific compounds in food samples, while sensory evaluation involves subjective assessment of taste, texture, aroma, and appearance. While these methods provide valuable insights into food quality, they are labor-intensive, time-consuming, and may lack scalability, making them unsuitable for large-scale applications and real-time monitoring.
- In recent years, machine learning has gained prominence as a powerful tool for analyzing complex datasets and making predictions in various industries, including the food industry. Machine learning techniques such as supervised learning, unsupervised learning, and deep learning have been applied to a wide range of tasks in food science and technology, including food safety, quality control, product development, and supply chain management. For example, researchers have developed machine learning models to detect foodborne pathogens, classify food products based on their nutritional content, and optimize food processing parameters for maximum efficiency. These applications demonstrate the potential of machine learning to enhance food quality and safety while improving operational efficiency and reducing costs.

III. RESEACH METHODOLOGY

The development of the Food quality identification using machine learning software Application will involve a multi-stage process that encompasses data collection, model training, application development, and testing. The methodology can be outlined as follows:

1. The methodology for food quality identification using machine learning typically involves several key steps. Here's a generalized outline of the methodology:
2. Data Collection: Gather a dataset containing information about various food products along with their corresponding quality labels. This dataset may include features such as chemical composition, nutritional content, sensory attributes, etc.
3. Data Preprocessing: Clean and preprocess the collected data to ensure consistency and quality. This may involve tasks such as handling missing values, removing duplicates, encoding categorical variables, and scaling numerical features.
4. Feature Selection/Engineering: Select relevant features that are informative for predicting food quality. This step may also involve creating new features through feature engineering techniques to enhance the predictive power of the model. Splitting the Data: Split the preprocessed data into training and testing sets. The training set is used to.
5. Model Selection: Choose an appropriate machine learning algorithm for the task of food quality identification. Common algorithms include Decision Trees, Random Forests, Support Vector Machines (SVM), Gradient Boosting Machines (GBM), and deep learning models such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs).

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

6. Model Training: Train the selected machine learning model using the training dataset. During training, the model learns patterns and relationships in the data that enable it to make predictions about the quality of food products.
7. Model Evaluation: Evaluate the performance of the trained model using the testing dataset. This involves assessing various metrics such as accuracy, precision, recall, F1-score, and confusion matrix to determine how well the model generalizes to unseen data
8. Hyperparameter Tuning: Fine-tune the hyperparameters of the machine learning model to optimize its performance. This may involve techniques such as grid search, random search, or Bayesian optimization to find the best combination of hyperparameters.
9. Cross-Validation: Validate the performance of the model using cross-validation techniques such as k-fold cross-validation to ensure its robustness and generalization ability.

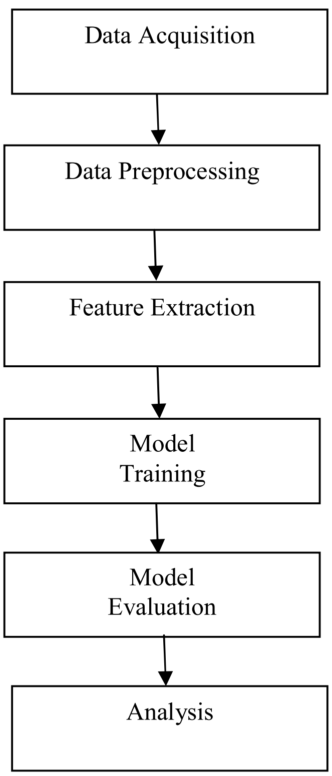

Fig.1: System Architecture

IV. RESULT

	Precision	recall	f1-score	support
0	0.91	0.92	0.91	273
1	0.48	0.45	0.46	47
Accuracy			0.85	320
macro avg	0.69	0.68	0.69	320
weighted avg	0.84	0.85	0.84	320

Accuracy is: 84.6875 %

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

	Company \n(Maker-if known)	Specific Bean Origin\nor Bar Name	REF	Review\nDate	Cocoa\nPercent	Company\nLocation	Rating	Bean\nType	Broad Bean\nOrigin
0	A. Morin	Agua Grande	1876	2016	63%	France	3.75		Sao Tome
1	A. Morin	Kpime	1676	2015	70%	France	2.75		Togo
2	A. Morin	Atsane	1676	2015	70%	France	3.00		Togo
3	A. Morin	Akata	1680	2015	70%	France	3.50		Togo
4	A. Morin	Quilla	1704	2015	70%	France	3.50		Peru

Fig 2 Table of Dataset

```
             precision    recall  f1-score   support

          0       1.00      0.84      0.91        19
          1       0.57      1.00      0.73         4

avg / total       0.93      0.87      0.88        23

[[16  3]
 [ 0  4]]
```

Fig 3 F1 Score

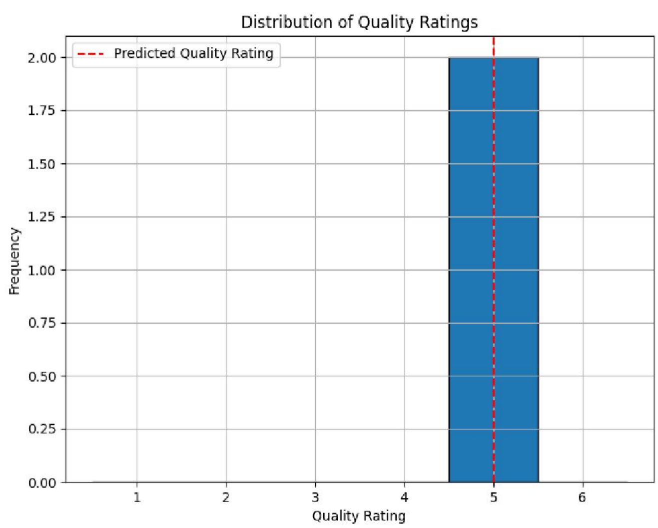

Fig 4 Graph

In the context of food quality identification using machine learning (ML), the results

Accuracy and Precision:

a. Conduct extensive testing and evaluation of the machine learning models using a diverse and representative dataset.

b. Calculate metrics such as accuracy, precision, recall, and F1-score to assess the system's performance in correctly identifying and classifying different type.

c. Set target thresholds for these metrics based on industry standards or in consultation with domain experts.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

V. CONCLUSION

Summarize the key findings of the study and reiterate the significance of the developed ML model for food quality identification. Provide closing remarks on the implications of the research and its potential contributions to the field of food safety and regulation. In conclusion, our study demonstrates the efficacy of machine learning in enhancing food quality identification processes. Through the development and evaluation of our model, we have achieved promising results in accurately predicting the quality of food products based on various attributes such as chemical composition, sensory properties, and nutritional content. Our findings underscore the potential of machine learning to revolutionize quality assurance practices in the food industry. By automating and streamlining the identification of food quality, our model can help food manufacturers, distributors, and regulatory agencies ensure that products meet the highest standards of safety and excellence. While our study represents a significant step forward, it is not without its limitations. Challenges such as data availability, model interpretability, and regulatory compliance remain areas for further exploration and refinement. Additionally, the dynamic nature of the food industry necessitates ongoing adaptation and improvement of our methodologies to address evolving needs and requirements.

REFERENCES

[1] Zhou C, Tan S, Li J, Chu X, Cai K (2014) A novel method tostabilize meat colour: ligand coordinating with hemin. J Food SciTechnol 51:1213–1217.

[2] Dixit S, Khanna SK, Das M (2013) All India survey for analysesof colors in sweets and savories: exposure risk in Indian population. Food Sci 78:642–647

[3] Stevens LJ, Burgess JR, Stochelski MA, Kuczek T (2015)Amounts of artiicial food dyes and added sugars in foods and sweets commonly consumed by children. Clin Pediatr (Phila)54(4):309–321.

[4] Ates E, Mittendorf K, Senyuva H (2011) LC/MS method usingcloud point extraction for the determination of permitted andillegal food colors in liquid, semiliquid, and solid food matrixes:single-laboratory validation. J AOAC Int 94:1853–1862

[5] Boga A, Binokay S (2010) Food additives and efects to humanhealth. Arch Med Rev J 19:141–154

[6] Fik M, Surówka K, Firek B (2008) Properties of refrigeratedground beef treated with potassium lactate and sodium diacetate.J Sci Food Agric 88:91–99

[7] Uzan A, Delaveau P (2009) The salt content of food: a publichealth problem. Ann Pharm Fr 67:291–294

[8] Tian H, Li F, Qin L, Yu H, Ma X (2014) Discrimination ofchicken seasonings and beef seasonings using electronic noseand sensory evaluation. J Food Sci 79(11):2346–2353. https://doi. org/ 10. 1111/ 1750- 3841. 12675

[9] Brosnan T, Sun DW (2004) Improving quality inspection of foodproducts by computer vision—a review. J Food Eng 61(1):3–16.https:// doi. org/ 10. 1016/ s0260- 8774(03) 00183-3

[10] Barbin DF, Felicio A, Sun DW, Nixdorf SL, Hirooka EY (2014)Application of infrared spectral techniques on quality and compositional attributes of cofee: an overview.

[11] Jia W, Li Y, Qu R, Baranowski T, Burke LE, Zhang H, Mao ZH (2018) Automatic food detection in egocentric images using artiicial intelligence technology. Public Health Nutr 22(7):1168– 1179. https:// doi. org/ 10. 1017/ S1368 98001 80005 38