**AI Nose for Food Spoilage Detection**



Prepared For

DR. Muhammad Farrukh Shahid  
Professor  
FAST university

By

21k-3062 Hairam Naseem

21k-3061 Haris Jamal

21k-3006 Mohib Ahmed

Bs Artificial Intelligence

FAST University

December 8, 2023

**Table of Contents**

i. Abstract ------------------------------------------------------------------------------------------------------------------4

1. Acknowledgements ---------------------------------------------------------------------------------------------------5

2. Introduction ------------------------------------------------------------------------------------------------------------5

3. Background -------------------------------------------------------------------------------------------------------------5

4. Description of Algorithms -------------------------------------------------------------------------------------------5

4.1 Artificial Neural Network (ANN) --------------------------------------------------------------------5

4.1.1 Principles of ANN -------------------------------------------------------------------------5

4.1.2 Application in Food Spoilage Detection ---------------------------------------------5

4.1.3 Advantages ---------------------------------------------------------------------------------5

4.2 k-Nearest Neighbors (KNN) --------------------------------------------------------------------------6

4.2.1 Principles of KNN -------------------------------------------------------------------------6

4.2.2 Application in Food Spoilage Detection ---------------------------------------------6

4.2.3 Advantages ---------------------------------------------------------------------------------6

4.3 Support Vector Machine (SVM) --------------------------------------------------------------------6

4.3.1 principles of SVM -------------------------------------------------------------------------6

4.3.2 Application in Food Spoilage Detection ---------------------------------------------6

4.3.3 Advantages ---------------------------------------------------------------------------------6

4.4 Logistic Regression ------------------------------------------------------------------------------------7

4.4.1 Principles of Logistic Regression ------------------------------------------------------7

4.4.2 Application in Food Spoilage Detection ---------------------------------------------7

4.4.3 Advantages --------------------------------------------------------------------------------7

5. Implementation -------------------------------------------------------------------------------------------------------7

6. Experiment Design ---------------------------------------------------------------------------------------------------8

6.1 Objective ------------------------------------------------------------------------------------------------8

6.2 Hypothesis ----------------------------------------------------------------------------------------------8

6.3 Limitations ----------------------------------------------------------------------------------------------8

7. Algorithm Comparison Results ------------------------------------------------------------------------------------9

8. Future Work ----------------------------------------------------------------------------------------------------------10

9. Conclusion ------------------------------------------------------------------------------------------------------------10

10. Reflection on Learning ---------------------------------------------------------------------------------------------10

11. References ------------------------------------------------------------------------------------------------------------10

12. Glossary ---------------------------------------------------------------------------------------------------------------11

13. Table of Abbreviations ---------------------------------------------------------------------------------------------11

Abstract

Ensuring safe food supplies has recently become a serious problem all over the world. Controlling the quality, spoilage, and standing time for products with a short shelf life is a quite difficult problem. However, electronic noses can make all these controls possible. In this study, which aims to develop a different approach to the solution of this problem, electronic nose data obtained from 12 different beef cuts were classified. In the dataset, there are four classes (1: excellent, 2: good, 3: acceptable, and 4: spoiled) indicating beef quality. The classifications were performed separately for each cut and all cut shapes. The ANOVA method was used to determine the active features in the dataset with data for 12 features. The same classification processes were carried out by using the three active features selected by the ANOVA method. Three different machine learning methods, Artificial Neural Network, K Nearest Neighbor, and Logistic Regression, which are frequently used in the literature, were used in classifications. In the experimental studies, a classification accuracy of 98% was obtained as a result of the classification performed with ANN using the data obtained by combining all the tables in the dataset.

1. Acknowledgements

We would like to express our gratitude to Dr. Muhammad Farrukh Shahid for his guidance and support throughout the course of this research project.

1. Introduction

Food spoilage poses a significant threat to public health and safety. This report explores the use of artificial intelligence techniques to detect food spoilage at an early stage, with a focus on the implementation of different ML algorithms specifically on the Beef Dataset.

1. Background

In recent years, advancements in artificial intelligence have opened new avenues for enhancing food safety. The concept of an AI nose involves the use of machine learning algorithms to analyze the volatile compounds emitted by spoiling food, enabling early detection.

1. Description of Algorithms
   1. Artificial Neural Network (ANN)
      1. Principles of ANN

Artificial Neural Networks (ANNs) are computer models that are based on how the human brain is organized and functions. Artificial neural networks (ANNs) use volatile chemical data to identify complicated patterns that are used to forecast food spoilage.

* + 1. Applications in Food spoilage detection

ANNs' quick processing speeds make real-time monitoring possible, which is essential for taking quick judgments to stop eating bad food. Additionally, by precisely recognizing spoiling symptoms, ANNs are useful in quality control procedures, guaranteeing that only safe and fresh food goods reach consumers.

* + 1. Advantages

There are several benefits to using ANNs in this situation. Because ANNs are so good at pattern recognition, they can identify subtle variations in volatile component concentrations that could indicate that food is spoiling. Since they are adaptable, they can learn and adapt continuously, which makes them ideal for dynamic contexts where spoiling properties may change over time.

* 1. k-Nearest Neighbors (KNN)
     1. Principles of KNN

The k-Nearest Neighbors (KNN) algorithm is a simple, yet effective supervised machine learning method used for both classification and regression tasks. The core idea behind KNN is to predict the class or value of a data point based on the classes or values of its neighboring data points in the feature space.

* + 1. Applications in Food spoilage detection

In the context of food spoilage detection, KNN can be applied by representing each food sample as a point in a multidimensional space, where each dimension corresponds to a feature (e.g., volatile compounds). The distance between points in this space is indicative of their similarity or dissimilarity.

* + 1. Advantages

KNN is easy to understand and implement, making it a straightforward choice for tasks with clear spatial patterns.

KNN is non-parametric and adaptable to changes in the dataset, making it suitable for dynamic environments.

* 1. Support Vector Machine (SVM)
     1. Principles of SVM

In food spoilage detection, Support Vector Machines (SVMs) function by identifying the best hyperplane to maximally divide data points in a high-dimensional space. To differentiate between fresh and rotten food samples, SVMs use volatile compound data to draw decision boundaries in the context of food spoilage detection.

* + 1. Applications in Food spoilage detection

SVMs are useful in applications where they can generate optimal decision boundaries in high-dimensional feature spaces, which helps with precise food spoilage detection. They work well in situations where there are intricate interactions between volatile substances and spoiling because of their ability to handle complex and non-linear correlations.

* + 1. Advantages

SVMs work effectively in situations when there are many of features.  
SVMs also provide generalization features that guarantee dependable performance on untested data and are resistant to overfitting.

* 1. Logistic Regression
     1. Principles of Logistic Regression

In the field of food spoilage detection, logistic regression is used to model the likelihood that a given sample falls into a specific class, such as fresh or spoilt. The logistic function is used in this binary classification process to generate an output between 0 and 1, which represents the likelihood that a sample belongs to a particular class.

4.4.2 Applications in Food spoilage detection

By offering probability estimates for various classes, logistic regression helps detect food rotting in real-world situations. These estimates can be thresholded to produce binary forecasts, and the model is useful in situations where it is important to comprehend the underlying relationships because of its simplicity and interpretability.

* + 1. Advantages

In this situation, the benefits of using Logistic Regression include its ease of use, interpretability, and effectiveness in managing linear correlations between the target variable and features. When there is a mostly linear association between volatile chemicals and spoiling, logistic regression is a good fit.

1. Implementation

The first step in utilizing AI noses to identify food deterioration is to carefully choose a variety of food samples and use electronic noses or gas sensors to detect volatile substances. The gathered data in our case is of beef which is then carefully preprocessed, which includes cleaning, feature extraction, and normalization, to make sure it is appropriate for algorithmic analysis.

The algorithms that can handle multidimensional feature spaces include Random Forest (RF) and k-Nearest Neighbors (KNN). While the KNN model normalizes features and classifies data based on the majority class of k-nearest neighbors, the RF model uses feature randomization and is trained via a majority voting method.

For developing a model, the dataset is then divided into training and testing sets. Cross-validation is then used to ensure a robust evaluation. Performance measures including precision and accuracy are examined, and the outcomes of the RF and KNN models are contrasted.

This implementation technique offers a thorough process for developing an efficient artificial intelligence (AI) nose system for early food rotting detection.

1. Experiment Design

6.1 Objective

This experiment's main goal is to evaluate and compare the effectiveness of four machine learning algorithms in a classification task: K-Nearest Neighbors (KNN), Artificial Neural Network (ANN), Logistic Regression, and Support Vector Machine (SVM). Finding the algorithm that best satisfies the project's objectives in terms of recall, accuracy, precision, and specificity is the aim. The purpose of this evaluation is to help choose the best model for the given job and dataset by offering insightful information about the advantages and disadvantages of each algorithm.

6.2 Hypothesis

Null Hypothesis (H0):  
The null hypothesis for this experiment posits that there is no significant difference in the performance metrics (accuracy, precision, recall, specificity) among the four machine learning algorithms—K-Nearest Neighbors (KNN), Artificial Neural Network (ANN), Logistic Regression, and Support Vector Machine (SVM).

Alternative Hypothesis (H1):  
Conversely, the alternative hypothesis suggests that there is a statistically significant difference in the performance metrics among the evaluated algorithms. More specifically, it posits that at least one algorithm outperforms the others in terms of accuracy, precision, recall, or specificity.

6.3 Limitations

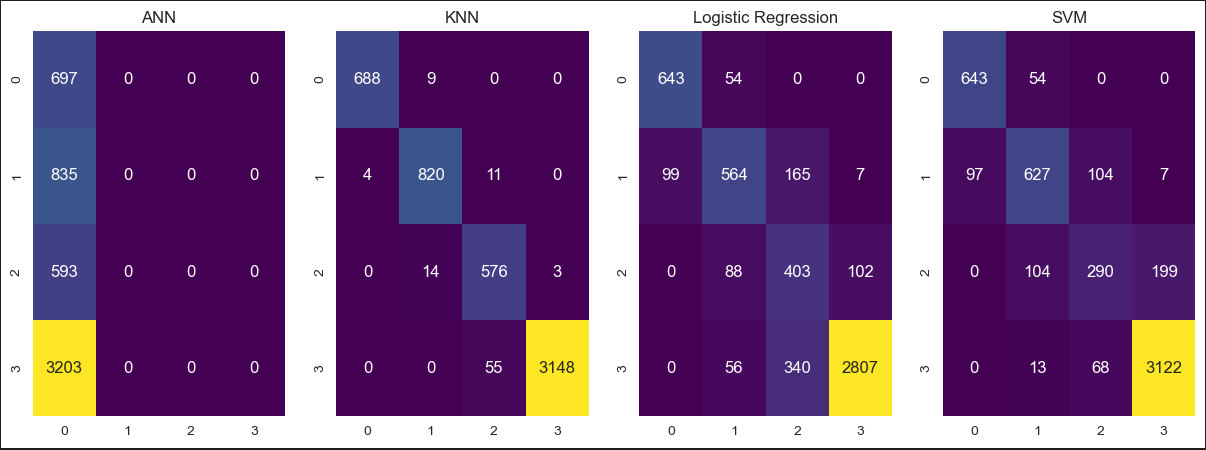
Firstly, the performance of machine learning algorithms is heavily dependent on the characteristics of the dataset. Variability in data distribution, size, and quality may influence the generalizability of the findings.

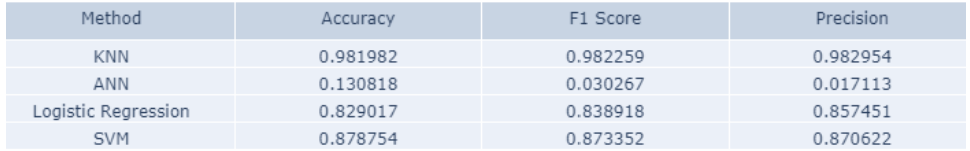
Additionally, the experiment assumes that the selected metrics (accuracy, precision, recall, specificity) adequately represent the desired criteria for algorithm evaluation.

Lastly, the results are subject to the specific hyperparameters and configurations used for each algorithm, and fine-tuning may be required for optimal performance in a real-world scenario.

1. Algorithm Comparison Results

For classification tasks in the research, four distinct machine learning algorithms were used, and their performance indicators were assessed. With excellent accuracy, precision, F1 score, recall, and overall performance, K-Nearest Neighbors (KNN) demonstrated better performance. It did, however, show comparatively less specificity when measured against other algorithms. The Artificial Neural Network (ANN) exhibited exceptionally low performance on all metrics, including a particularly low F1 score and accuracy, as well as an inability of the model to compute specificity (NaN). Respectable findings were obtained from the logistic regression, including good recall, accuracy, and precision as well as a comparatively high specificity. Among other metrics, Support Vector Machine (SVM) demonstrated good performance, making it a formidable competitor.





1. Future Work

* Sensor Improvement:  
  Investigate advanced sensor technologies for more accurate volatile compound detection.
* Integration with IoT:  
  Explore the integration of the AI nose with Internet of Things (IoT) devices for real-time monitoring.
* Extended Dataset:  
  Expand the dataset to include a wider variety of foods and spoilage conditions for a more robust model.

1. Conclusion

This experiment offers insightful information about algorithm selection for AI nose applications, highlighting the significance of customizing the selection to the particulars of the data and task at hand. The results highlight the need for continual development and improvement in the field of AI-based spoiling detection systems and contribute to the ongoing efforts to improve food safety using innovative technology.

1. Reflection on Learning

The research has been a fascinating investigation into the principles of machine learning and how they are applied in a particular field. The knowledge and abilities gained are priceless, providing the foundation for further development in the ever-expanding field of artificial intelligence and its numerous applications. In the ever-changing world of technology, the reflective approach emphasizes the value of curiosity, flexibility, and a dedication to lifelong learning.

1. References

* Wilson, A. D., & Baietto, M. (2009). Applications and advances in electronic-nose technologies. Sensors, 9(7), 5099-5148. doi:10.3390/s90705099
* Shokouhi, S., & Kermani, M. J. (2016). An electronic nose based on a micro-machined gas sensor array to detect spoiled food. Sensors and Actuators B: Chemical, 237, 749-757. doi:10.1016/j.snb.2016.06.040
* Persaud, K. C., & Dodd, G. (1982). Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose. Nature, 299(5880), 352-355. doi:10.1038/299352a0

1. Glossary

**AI Nose**: An artificial intelligence system designed to mimic the olfactory capabilities of the human nose. It involves the use of sensors to detect and analyze volatile compounds in the air, enabling applications in various fields, including food spoilage detection.

**Hyperparameter:** A configuration setting external to a model that cannot be learned from training data and must be set prior to training.

**Volatile Compounds**: Chemical compounds that can evaporate into the air, often responsible for the characteristic odor of substances.

**Internet of Things (IoT):** A network of interconnected physical devices that communicate and exchange data with each other, contributing to the automation and optimization of various processes.

1. Table of Abbreviations

|  |  |
| --- | --- |
| Abbreviation | Definition |
| AI | Artificial Intelligence |
| KNN | K-Nearest Neighbor |
| SVM | Support vector machine |
|  |  |