

Application of Machine Learning Techniques for Okra Shelf Life Prediction

Iveren Blessing Iorliam¹, Barnabas Achakpa Ikyo¹, Aamo Iorliam^{2*}, Emmanuel Odeh Okube², Kenneth Dekera Kwaghtyo², Yahaya I. Shehu³

¹Centre for Food Technology and Research (CEFTEER), Benue State University, Makurdi, Nigeria

²Department of Mathematics & Computer Science, BSU, Makurdi, Nigeria

³Shehu Shagari College of Education, Sokoto, Nigeria

Email: *aamoiorliam@gmail.com

How to cite this paper: Iorliam, I.B., Ikyo, B.A., Iorliam, A., Okube, E.O., Kwaghtyo, K.D. and Shehu, Y.I. (2021) Application of Machine Learning Techniques for Okra Shelf Life Prediction. *Journal of Data Analysis and Information Processing*, 9, 136-150.
<https://doi.org/10.4236/jdaip.2021.93009>

Received: May 20, 2021

Accepted: July 17, 2021

Published: July 20, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The ability of machine learning techniques to make accurate predictions is increasing. The aim of this work is to apply machine learning techniques such as Support Vector Machine, Naïve Bayes, Decision Tree, Logistic Regression, and K-Nearest Neighbour algorithms to predict the shelf life of Okra. Predicting the shelf life of Okra is important because Okra becomes harmful for human consumption if consumed after its shelf life. Okra parameters such as weight loss, firmness, Titrable Acid, Total Soluble Solids, Vitamin C/Ascorbic acid content, and PH were used as inputs into these machine learning techniques. Support Vector Machine, Naïve Bayes and Decision Tree each accurately predicted the shelf life of Okra with accuracies of 100%. However, the Logistic Regression and K-Nearest Neighbour achieved 88.89% and 88.33% accuracies, respectively. These results showed that machine learning techniques especially Support Vector Machine, Naïve Bayes and Decision Tree can be effectively applied for the prediction of Okra shelf life.

Keywords

Machine Learning, Shelf Life, Okra, Prediction

1. Introduction

Okra (*Hibiscus esculentus*) is considered as the second most important vegetable in West African market after tomatoes [1] [2].

The vegetable crop is known for its mucilaginous properties that give the Okra soup a slimy texture, which has huge benefits as stabilizers, emulsifiers and thickeners [3]. Okra is a versatile crop known for its tender and tasty pods as

well as its importance in preparing African soups and sauces. In Nigeria and most African regions, Okra soup is prepared from fresh or dried Okra pods.

Although, it is believed that the nutritional value of Okra when preserved is altered, research has it that such alteration is almost insignificant.

A comparison shows that the dried Okra (Okra chips) retains the mucilage from the fresh Okra pods. Okra is considered as an important fruit because it contains carbohydrates, proteins, total fat, cholesterol, dietary fibre, vitamins, electrolytes, minerals and phyto-nutrients [4]. Furthermore, Okra has health benefits such as high fiber content, prevention of diabetes, serves as good folate content, provides Vitamin K benefits, controls asthma, prevents constipation, prevents sunstroke, prevents colon cancer, controls obesity, controls cholesterol level, has low Glycemic Index, prevents kidney diseases, helps in digestion, good source of antioxidants, prevents skin pigmentation, improves immunity, improves eye sight, and supports ulcer healing amongst several other benefits [5]. With the huge health benefits that Okra provides, it is very important to extend its shelf life.

Most of the farmers/traders use synthetic chemicals such as actellic 25 EC (pirimiphos-methyl), heptachlor, and thiometon to extend the shelf life of Okra [6]. These synthetic chemicals are harmful to the environment and their chemical residues may be found in the Okra fruit. The chemicals are also dangerous to the health of the users [7]. It is worth noting that biological approaches in preserving general vegetables have proved to be very effective and suitable for mammals [8]. Even though there are methods developed to extend Okra shelf life, not much has been done in predicting the shelf life of Okra.

For the fact that machine learning techniques have the capability to learn from existing data and make predictions, this research looks at how safe Okra is for consumption either in the preserved or unpreserved state using machine learning techniques. By preserved state, we mean Okra that is treated to extend its shelf life, whereas, unpreserved state means Okra that is harvested and not treated.

In this paper, we propose the novel use of support vector machine (SVM), Naïve Bayes, Decision Tree, Logistic Regression and K-Nearest Neighbour (KNN) to predict how long Okra can stay (shelf life) before it is no longer safe for consumption based on some measured Okra parameters.

2. Related Work

Research attempts in predicting the shelf life of fruits are generally rare. However, some attempts have been made by some researchers for the classification of fruits.

Al-Sulaiman [9], developed a framework for predicting the quality of Okra pods after drying in a domestic microwave oven using an artificial neural network (ANN) model. The evaluation results indicated that the ANN technique outperformed the multiple linear regression model for the prediction of Okra

quality. The model achieved an R^2 value of 0.98.

Savakar [10], proposed a classification technique based on artificial neural network (ANN) to identify the diverse types of fruits. Five different types of fruits, *i.e.*, Apple, Mango, Sweet Lemon, Chikoo and Orange were considered and a total of 5000 images were used. Each of the fruit types contributed 1000 images that were examined using the ANN model. Experiments demonstrated that the model performed well with an accuracy of 94%.

Dubey and Jalal [11], presented a model for fruit and vegetable classification using the K-Means clustering method for segmentation. The classification of fruits and vegetables was done by multi-class SVM, based on the improved sum and difference histogram (ISADH) texture feature, which was a modified version of sum and difference histograms for texture classification proposed by Unser in 1986. Results showed that the model recorded an accuracy of 99%.

Yissah, Ikyo and Ige [12], proposed the extension of Okra shelf-life using X-ray irradiation of varying quantity. The technique elongates the shelf-life of fresh Okra to minimize its post-harvest loss and increase its availability. Experiments using an irradiation dosage of 0.051 Gy as recommended by NAFDAC and USFD, demonstrated that the approach prolonged the fresh Okra shelf-life from 3 days to 14 days.

Yu, Li, Shen, Cao and Mao [13], employed the use of Bias classifier, decision tree and SVM techniques to classify and compare their results to determine the most efficient algorithm in the prediction of grain data. Experiments showed that the SVM algorithm outperformed the other algorithms.

Khaing, Naung, and Htut [14], developed a control system of objection recognition based on a convolutional neural network (CNN). The CNN was given the task of fruits detection and recognition through parameter optimization. The result of the model had an accuracy of 94% for the 30 classes of 971 images. This indicated that the proposed method could be used for vision subsystem based control applications.

Golpour, Chayjan, Parian and Khazaei [15], developed a machine learning model for the prediction of the moisture content of paddy using artificial neural network (ANN). The model was subjected to six different temperatures and five velocities for experimental purpose. Experimental results demonstrated that the model achieved a mean square error of 0.001, mean absolute error of 0.03 and R^2 value of 0.9630 for the prediction.

Sakib, Ashrafi, Siddique and Bakr [16], proposed a deep learning model using convolutional neural network (CNN) for fruit identification or recognition. The fruits-360 dataset was used for testing the proposed model. Results showed that the model achieved an overall accuracy of 100% and a training accuracy of 99.79%.

Nosseir, and Ahmed [17], developed a model that classified four different types of fruits and further separated the decayed ones from the fresh ones. The model adopted the use of K-NN and SVM techniques for the prediction. Experiments demonstrated that the model achieved 95%, 96% and 98% accuracies

using K-NN, linear SVM and quadratic SVM, respectively.

Ummapure and Hanchinal [18], proposed a multi-class framework where three feature vectors were developed to identify fruits types. The multi-class framework achieved 99.98% classification accuracy.

Zeeshan, Prabhu, Arun and Rani [19], implemented a fruit classification system using computer vision and support vector machine (SVM). The model employed the use of the Gaussian filter which reduced the noise while enhancing the image quality for effective classification. The model was tested using 655 images and achieved an overall accuracy of 87.06%.

Behera, Rath, Mahapatra and Sethy [20], presented a comprehensive analysis of the existing methods for fruits classification with a focus on the development of state-of-the-art methods. The study compared different techniques for the identification, classification and grading of fruits.

Jaramillo-Acevedo, Choque-Valderrama, Guerrero-Álvarez and Meneses-Escobar [21], leveraged on the RGB color model which is based on the physical and chemical changes during the ripening process. Hence, classified the consumption maturity for Hass avocado fruits using an artificial neural network. Experimental results yielded 88% ripeness estimate accuracy and 0.819 regression value during the post-harvest period.

Ogbaji and Iorliam [22], carried out a laboratory experiment and showed that the shelf life of the treated Okra fruits ranged from days 1 - 15, while the control ranged from 1 - 7 days. The laboratory data generated from their research is used in this paper for the shelf life prediction of Okra.

To the best of the researchers' knowledge, the prediction of Okra shelf life has not been investigated using machine learning techniques. Hence, this research is motivated to apply machine learning techniques to predict Okra shelf life.

3. Materials and Methods

3.1. Methodology

In this paper, five algorithms including SVM, Naïve Bayes, Decision Tree, Logistic Regression and KNN are employed to predict and compare the shelf life of Okra. Three scenarios of:

- 1) Okra without preservation;
- 2) Okra preserved with Moringa and
- 3) Okra preserved with Neem was considered.

Data generated from the three scenarios were fed into the five algorithms as input. In each case, the algorithms did a prediction to determine whether the Okra is safe for consumption or not based on the laboratory data generated. The performances of all the five algorithms are presented in terms of the Accuracy, F1-score, Recall, and Precision.

3.2. Dataset Used

The measurement of the Okra fruits weight loss (%) was achieved by placing the

fruits on a digital weighing balance and each reading was recorded during the storage process.

The fruit firmness (N/cm) was measured as the maximum penetration force (N) reached during tissue breakage using a standard probe. The registered force at the penetration of a standard probe up to a certain depth (cm) was read as firmness. The firmness of the Okra fruit was determined using a penetrometer as described in Kumah, Olympio, and Tayviah [23].

The percentage titratable acidity (TA %) of Okra fruits was determined as follows; after blending the Okra fruits, Ten millilitres of the juice were filtered using a filter paper in a beaker. Five millilitres of the filtrate were pipetted into a conical flask, then 10 millilitres of sterile distilled water were added to aid a clear endpoint detection. After this, two drops of phenolphthalein indicator were added. 0.1 N sodium hydroxide (NaOH) was added dropwise and the solution was shaken thoroughly until a pink colour was obtained. The acid content of the Okra sample was calculated using the formula:

$$\% \text{ TA} = \frac{V * M * F * 100}{\text{volume of Okra juice}}$$

where; V = volume of 0.1 N NaOH used,

M = Molarity of NaOH and,

F = factor of citric acid (0.0064).

The Total Soluble Solids (TSS) (°Brix) of the Okra fruits was determined using a handheld refractometer. A homogenous sample was obtained after blending Okra fruits, then two drops from the blended sample were applied to the refractometer using a plastic dropper, and the reading was obtained as a percentage of soluble solids concentration in °Brix [24].

Vitamin C/Ascorbic acid content (mg/ml) was determined using the method described in AOAC [24].

$$\text{Molarity of indophenol} = \frac{\text{Contraction of Volume C} * \text{Volume of Vitamin}}{\text{Volume of indophenol}}$$

The final volume was recorded and the concentration of Vitamin C in the sample was calculated and expressed in mg/ml using the above formula. Okra fruits pH was determined by inserting the pH meter into the Okra paste obtained after blending and the readings obtained [25].

The shelf life of Okra was evaluated by counting the number of days Okra fruits were still acceptable for marketing and consumption. This was decided based on the appearance and spoilage of the Okra fruits. The above process is reported vividly in Ogbaji and Iorliam [22].

3.3. Data Preprocessing

Data pre-processing in this study was the removal of noise from experimental data by sieving out unwanted parameters. The unprocessed data was pre-processed to align with core parameters such as Firmness, TSS, TTA, Vitamin C and Weight

loss of the Okra to predict the shelf of Okra as shown in **Figure 2**.

For the normalisation of data (which involves scaling down the dataset to eliminate the effects of variation in the datasets), the researchers obtained the normalization equation by subtracting the minimum value (X minimum) from the Okra variable to be normalized (X). After subtracting the minimum value (X minimum) from the maximum value (X maximum), the previous result is divided by the latter. Mathematically, normalisation was carried out using the equation below:

$$X \text{ normalised} = \frac{X - X \text{ minimum}}{X \text{ maximum} - X \text{ minimum}} \quad (1)$$

Variables like Treatment, Variety, Firmness, TSS, TTA, Vitamin C, and Weight loss were rearranged from the complex raw format to a less complex format (data cube aggregation). The resulting dataset became smaller in size (dimensionality reduction) while retaining all of the details required for prediction.

3.4. The Proposed Model

The designed model for the study is achieved by applying five different algorithms to make predictions from the dataset. Two classes are considered “class 1” and “class 0”. Class 1 is the class of Okra that is safe for consumption while class 0 is the class of Okra that is not safe for consumption. Taking into account some Okra parameters like Firmness, TSS, TTA, Vitamin C and Weight loss, the algorithms make independent predictions to determine the Precision values, Recall values, F1 values and Accuracies. For all the algorithms compared, the dataset is split into 70% training, and 30% testing. The proposed Okra shelf life prediction model is shown in **Figure 1**.

3.5. Model Description and Evaluation Metrics

3.5.1. SVM

Considering a function line $y = ax + by$. The support vector machine algorithm has a vector length denoted by x and the direction of the x represented by w . If

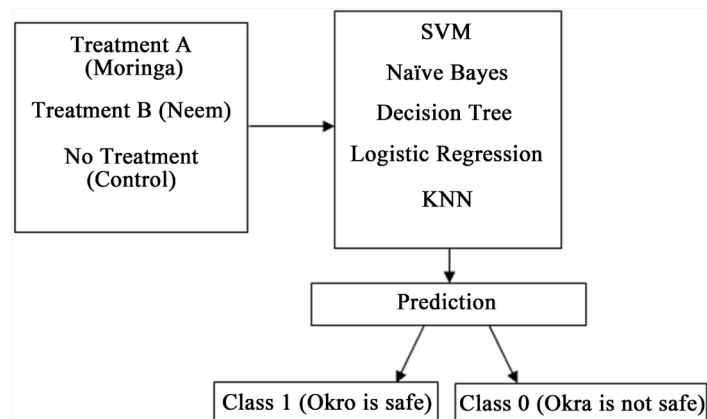


Figure 1. Proposed okra shelf life prediction model.

we define $x = (x_1, x_2)$ and $w = (a, -1)$,

We can get $w \cdot x + b = 0$.

The SVM classifier works with a hypothesis function h defined as:

$$h(x_i) = \begin{cases} +1, & \text{if } w \cdot x + b \geq 0 \\ -1, & \text{if } w \cdot x + b < 0 \end{cases} \quad (2)$$

Equation (2) is the mathematical representation of the support vector machine algorithm used in this study.

3.5.2. Naïve Bayes

Given a feature vector $X = (x_1, x_2, \dots, x_n)$ and a class variable C_k , Bayes Theorem states that:

$$P(C_k | X) = P(X | C_k)P(C_k)/P(X), \text{ for } k = 1, 2, \dots, K$$

We call $P(C_k | X)$ the posterior probability, $P(X | C_k)$ the likelihood, $P(C_k)$ the prior probability of a class, and $P(X)$ the prior probability of the predictor. In this study, we are concerned with calculating the posterior probability from the likelihood and prior probabilities.

Using the chain rule, the likelihood $P(X | C_k)$ can be decomposed as:

$$P(c | x) = \frac{P(x | c)P(c)}{P(x)}$$

$$P(c | x) = P(x_1 | c) * P(x_2 | c) * \dots * P(x_n | c) * P(c) \quad (3)$$

where;

$P(c|x)$ is the posterior probability of *class* (c , *target*) given *predictor* (x , *attributes*).

$P(c)$ is the prior probability of *class*.

$P(x/c)$ is the likelihood which is the probability of the predictor given class.

$P(x)$ is the prior probability of the predictor.

3.5.3. Decision Tree

The decision tree goes from observations about an item (represented in the branches) to conclusions about the item's target value. This algorithm uses two popular techniques for attribute selection: Information Gain and Gini Index. The Information Gain is given by:

$$\text{Information Gain} = \text{Entropy}(S) - [(\text{Weighted Avg}) * \text{Entropy}(\text{each feature})] \quad (4)$$

Entropy as a measure of impurity in a given attribute specifies randomness in data. Entropy is calculated as:

$$\text{Entropy}(s) = -P(\text{yes})\log 2P(\text{yes}) - P(\text{no})\log 2P(\text{no}) \quad (5)$$

where; s = total number of samples.

$P(\text{yes})$ = probability of yes.

$P(\text{no})$ = probability of no.

The Gini index which is a metric of impurity or purity used in the CART (Classification and Regression Tree) algorithm to build a decision tree is repre-

sented mathematically as:

$$\text{Gini Index} = 1 - \sum_j P_j^2 \quad (6)$$

where; j is the number of classes.

If a decision tree is used for regression and the output is continuous in nature, a reduction in variance is often used. Using the variance formula, the algorithm divides the population. The Reduction variance is expressed as

$$\text{Variance} = \frac{\sum (X - \bar{x})^2}{n}$$

where; \bar{x} is the mean of values, X is the actual mean and n is the number of values.

3.5.4. Logistic Regression

Logistic regression models the probability of an outcome depending on individual characteristics. It is given by:

$$\log(P_i/1 - P_i) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \quad (7)$$

where: P_i indicates the Bernoulli distribution of the variable under consideration (e.g., Okra shelf-life), β_i are the regression coefficients and the x_i are variables. The reference group (β_0) is the reference level of each and every variable x_{i1}, \dots, x_{ip} .

3.5.5. K-Nearest Neighbour (KNN)

KNN determines the closest point to the input point from the given data set. It thus stores the available cases (test data) and classifies new cases by majority votes of its K neighbours. The algorithm is based on which points share the highest probabilities using mathematical values (vectors). Hence, uses the Euclidean distance formula:

$$\text{dist}((x, y), (a, b)) = \sqrt{(x - a)^2 + (y - b)^2} \quad (8)$$

where (x, y) and (a, b) are coordinates.

3.5.6. Evaluation Metrics

The evaluation metrics used in this paper are the Accuracy, F1-score, Precision, Recall and Confusion Matrix.

4. Results and Discussion

This study used the dataset generated in the laboratory in Ogbaji and Iorliam [22], using the procedure explained in section 3.2. It consists of 60 rows and 8 columns of data stored in .XLSX (an Excel file format). It is a combination of both text and numbers as shown in **Figure 2**.

4.1. Support Vector Machine (SVM)

The performance accuracy of the SVM algorithm as shown in **Table 1** recorded a precision of 1.00, recall value of 1.0, F1 score of 1.00, and 100% accuracy. This

result affirms that the SVM algorithm performs well and can be relied upon for the prediction of Okra shelf life.

The confusion matrix for the SVM model is shown in **Figure 3**.

	A	B	C	D	E	F	G	H
1	treatment	variety	firmness	tss	tta	vit_c	weight_lc	shelf_life
2	Moringa	Var A	2.2	1.4	0.188	14.81	68.02	1
3	Moringa	Var A	2.2	1.5	0.187	14.964	64.54	1
4	Moringa	Var B	2	1.8	0.145	18.78	67.56	1
5	Moringa	Var B	2	1.9	0.142	18.78	67.04	1
6	Neem	Var A	2	1.4	0.206	14.76	73.33	1
7	Neem	Var A	2.1	1.4	0.211	14.81	73.2	1
8	Neem	Var B	1.7	1.6	0.158	18.724	72.32	1
9	Neem	Var B	1.8	1.5	0.17	18.58	77.4	1
10	Control	Var A	1.9	1.4	0.226	14.503	75.86	1
11	Control	Var A	2	1.5	0.218	14.611	75.55	1
12	Control	Var B	1.7	1.5	0.182	18.295	75.93	1
13	Control	Var B	1.7	1.5	0.169	18.4	78.71	1
14	Moringa	Var A	1.6	1	0.237	14.302	52.33	1
15	Moringa	Var A	1.7	1	0.25	14.14	55.68	1
16	Moringa	Var B	1.4	1	0.192	17.913	56.74	1
17	Moringa	Var B	1.4	1	0.192	18.105	52.27	1
18	Neem	Var A	1.5	1	0.293	14.105	58.69	1
19	Neem	Var A	1.4	1	0.28	13.94	60.51	1
20	Neem	Var B	1.3	1	0.215	17.967	59.6	1
21	Neem	Var B	1.1	1	0.228	17.704	59.14	1

Figure 2. Sample of the dataset used.

Table 1. Performance accuracy of the SVM algorithm.

Precision	Recall	F1 score	Accuracy
1.00	1.00	1.00	100.00%

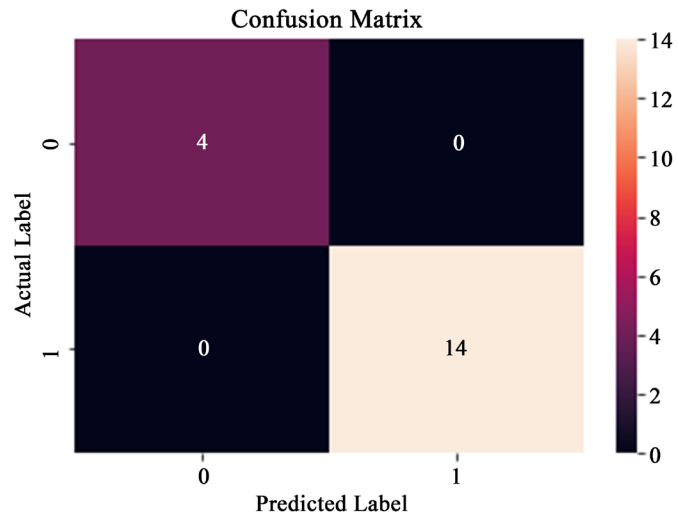


Figure 3. Confusion matrix for the SVM algorithm.

4.2. Naïve Bayes (NB)

From **Table 2**, the performance accuracy of the Naïve Bayes model also achieved 100% accuracy. The precision, recall values and F1 score all achieved 1.00. This shows that for the prediction of the shelf life of Okra, the Naïve Bayes algorithm also gave a very high prediction accuracy.

Figure 4 describes the performance accuracy of the Naïve Bayes model using a confusion matrix.

4.3. Decision Tree (DT)

The same dataset fed into the SVM and Naïve Bayes model was used in the Decision Tree model. **Table 3** shows the accuracy achieved in which the model achieved a Precision value of 1.00, recall value of 1.00, F1 score of 1.00, and accuracy of 100%.

The performance accuracy of the Decision Tree model is as shown in the confusion matrix in **Figure 5**.

4.4. Logistic Regression (LR)

Table 4 shows that the Logistic Regression algorithm achieved a Precision value of 0.88, recall value of 1, F1-score value of 0.80 and accuracy of 88.89%.

Using the Logistic Regression algorithm, the model yielded the confusion

Table 2. Performance accuracy of the naïve Bayes algorithm.

Precision	Recall	F1 score	Accuracy
1.00	1.00	1.00	100.00%

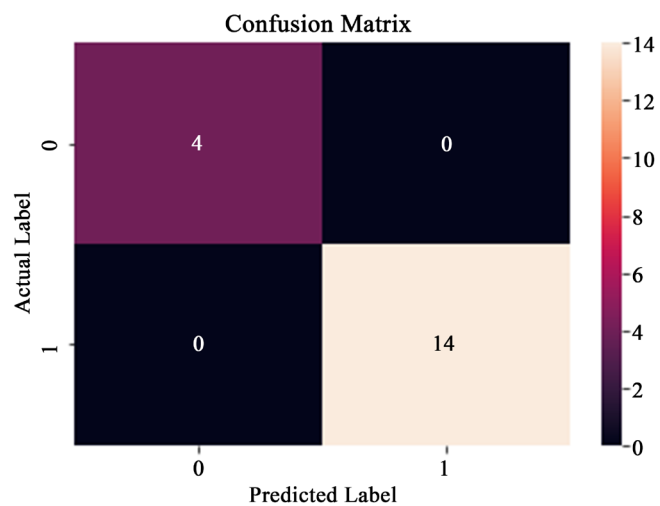


Figure 4. Confusion matrix for the naïve Bayes algorithm.

Table 3. Performance accuracy of the decision tree algorithm.

Precision	Recall	F1 score	Accuracy
1.00	1.00	1.00	100.00%

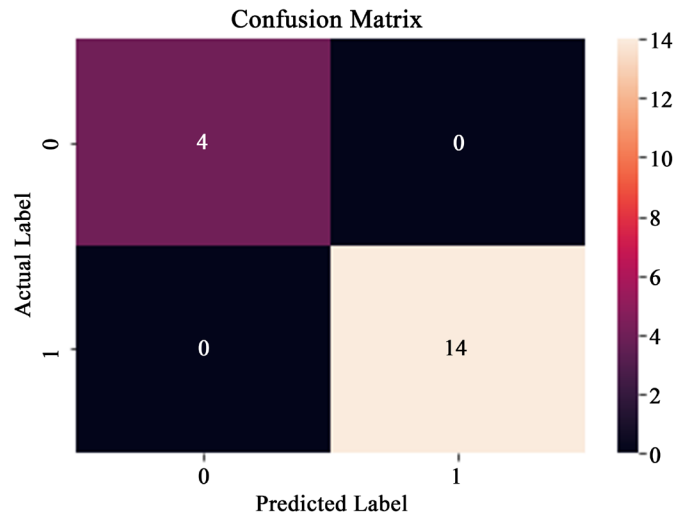


Figure 5. Confusion matrix for the Decision Tree algorithm.

Table 4. Performance accuracy of the logistic regression algorithm.

Precision	Recall	F1 score	Accuracy
0.88	1.00	0.80	88.89%

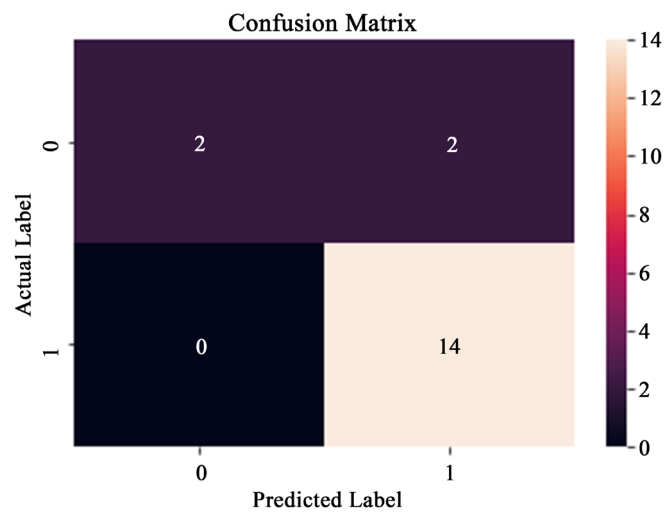


Figure 6. Confusion matrix for the logistic regression algorithm.

matrix as shown in Figure 6.

4.5. K-Nearest Neighbour (K-NN)

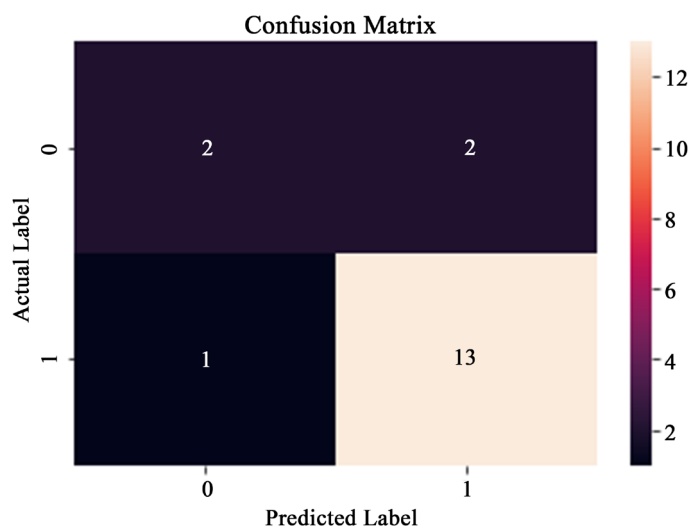
The KNN algorithm for predicting Okra shelf-life under study showed the Precision value of 0.87, recall value of 0.93, F1-score of 0.73, and accuracy of 83.33% (See Table 5).

The confusion matrix for the K-NN model is as shown in Figure 7.

Comparatively, Table 6 shows how our proposed model outperforms the works done by other researchers based on the accuracy achieved by the various models and approaches used. It is obvious from the comparison that our proposed

Table 5. Performance accuracy of the KNN algorithm.

Precision	Recall	F1 score	Accuracy
0.87	0.93	0.73	83.33%

**Figure 7.** Confusion Matrix for the K-NN Algorithm.**Table 6.** Comparative analysis with existing methods.

Reference	Title	Fruit(s)	Algorithm(s)	Accuracy(ies)
Dubey, and Jalal [11]	A robust approach for fruit and vegetable classification	Plum, potato, cashew, onion, orange, kiwi, apple, watermelon, pear and peach	SVM	99%
Savakar [10]	Identification and classification of bulk fruits images	Apple, mango, sweet lemon and orange	ANN	94%
Khaing, Naung, and Htut [14]	Control system for fruit classification	Apples, bananas, blueberries, kiwi fruits, raspberries	CNN	94%
Nosseir and Ahmed [17]	Automatic classification for fruits' types and identification of rotten ones	Banana, apple, and strawberry	K-NN, Linear SVM and Quadratic SVM	95%, 96% and 98%
Zeeshan, Prabhu, Arun and Rani [19]	Fruit classification system	Apple, Grapes, Bosch Pears, Banana, Blackberries, Blueberries, Anjou Pears, Cantaloupes, etc.	SVM	87 %
Ummature and Hanchinal [18]	Multi features based fruit classification	Apple, Banana, Cherry, Grapes, and Mango	SVM, MLP and RF	99.98%
Our Proposed Method	Application of machine learning techniques for Okra shelf life prediction	Okra fruits only	SVM, Naïve Bayes and Decision Tree	100%
			Logistic Regression	88.89%
			K-NN	83.33%

model outperformed the existing state-of-the-art methods especially when using the SVM, Naïve Bayes and Decision Tree. This proves that our proposed method is effective, efficient and can be relied upon for the prediction of Okra shelf life.

From **Figures 3-5**, it can be observed that the three algorithms (SVM, Naïve Bayes and Decision Tree), all predicted the shelf life of Okra correctly up to 100% accuracy. The confusion matrices show that in every case, 4 of the Okra are predicted to be bad (unsafe for consumption) and 14 of the Okra are predicted to be good (safe for consumption).

The low performance of the Logistic Regression and KNN algorithms shown in **Figure 6** and **Figure 7** further affirms the suitability of the SVM, Naïve Bayes and Decision Tree algorithms for the shelf-life prediction of Okra.

5. Conclusions and Future Work

In this work, we used five machine learning techniques (Support Vector Machine, Naïve Bayes, Decision Tree, Logistic Regression, and K-Nearest Neighbour) to predict the shelf life of Okra. Our prediction achieved an accuracy of 100% in the three algorithms (Support Vector Machine, Naïve Bayes and Decision Tree) and an F1 measure of 1.00, and a recall value of 1.00 for the three algorithms (Support Vector Machine, Naïve Bayes and Decision Tree). The Logistic Regression and KNN algorithms achieved lower prediction results. This shows that machine learning techniques are reliable and can accurately predict the shelf life of Okra.

In our future work, we hope to predict the shelf life of different variety of fruits using other machine learning techniques.

Acknowledgements

Iveren Blessing Iorliam sincerely thanks the management of Centre for Food Technology and Research (CEFTER), Benue State University Makurdi, Nigeria for funding the research that generated data used in this paper.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Ubani, O.N. and Okonkwo, E.U. (2011) A Review of Shelf-Life Extension Studies of Nigerian Indigenous Fresh Fruits and Vegetables in the Nigerian Stored Products Research Institute. *African Journal of Plant Science*, **5**, 537-546.
- [2] Kelechi, A.I., Jackson, N.N., Ifeyinwa, A.A. and Ajuruchukwu, O. (2013) Market Structure, Conduct, Channel and Margin of Dry Season Okra Vegetable in South-Eastern Nigeria. Department of Agricultural Economics, University of Nigeria, Nsukka.
- [3] Agbenorhevi, J.K., Kpodo, F.M., Banful, B.K., Oduro, I.N., Abe-Inge, V., Datsomor, D.N., Atongo, J. and Obeng, B. (2020) Survey and Evaluation of Okra Pectin Extracted at Different Maturity Stages. *Cogent Food & Agriculture*, **6**, Article ID: 1760476. <https://doi.org/10.1080/23311932.2020.1760476>
- [4] Agriexchange (2018) Okra.

- <http://apeda.in/agriexchange/Market%20Profile/one/OKRA.aspx>
- [5] Stylecraze (2018) 24 Amazing Benefits of Okra/Ladys Finger (Bhindi) for Skin, Hair, and Health. <http://www.factfish.com/statistic-country/nigeria/okra%2C%20production%20quantity>
- [6] Farinde, A.J., Owolarafe, O.K. and Ogungbemi, O.I. (2007) An Overview of Production, Processing, Marketing and Utilisation of Okra in Egbedore Local Government Area of Osun State, Nigeria. *Journal of Agricultural Engineering International*, **8**, 1-17.
- [7] Obeng-Ofori, D. and Sackey, J. (2003) Field Evaluation of Non-Synthetic Insecticides for the Management of Insect Pests of Okra (*Abelmoschus esculentus* (L.) Moench) in Ghana. *SINET: Ethiopian Journal of Science*, **26**, 145-150. <https://doi.org/10.4314/sinet.v26i2.18210>
- [8] Anjum, M.A., Ahmed, N., Babita, C.H. and Gupta, P. (2016) Plant Extracts in Post-Harvest Disease Management of Fruits and Vegetables—A Review. *Journal of Food Processing and Technology*, **7**, 592. <https://doi.org/10.4172/2157-7110.1000592>
- [9] Al-Sulaiman, M.A. (2011) Prediction of Quality Indices during Drying of Okra Pods in a Domestic Microwave Oven Using Artificial Neural Network Model. *African Journal of Agricultural Research*, **6**, 2680-2691.
- [10] Savakar, D. (2012) Identification and Classification of Bulk Fruits Images Using Artificial Neural Networks. *International Journal of Engineering and Innovative Technology*, **1**, 35-40.
- [11] Dubey, S.R. and Jalal, A.S. (2012) Robust Approach for Fruit and Vegetable Classification. *Procedia Engineering*, **38**, 3449-3453. <https://doi.org/10.1016/j.proeng.2012.06.398>
- [12] Yissah, S.O., Ikyo, B.A. and Ige, T.A. (2016) Extension of Fresh Okra Shelf-Life by X-Ray Irradiation. *Nigerian Journal of Pure and Applied Sciences*, **8**, 265-271.
- [13] Yu, H., Li, B., Shen, D., Cao, J. and Mao, B. (2017) Study on Prediction Model of Grain Post-Harvest Loss. *Procedia Computer Science*, **122**, 122-129. <https://doi.org/10.1016/j.procs.2017.11.350>
- [14] Khaing, Z.M., Naung, Y. and Htut, P.H. (2018) Development of Control System for Fruit Classification Based on Convolutional Neural Network. 2018 *IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering*, St. Petersburg, 29 January-1 February 2018, 1805-1807. <https://doi.org/10.1109/EIConRus.2018.8317456>
- [15] Golpour, I., Chayjan, R.A., Parian, J.A. and Khazaei, J. (2018) Prediction of Paddy Moisture Content during Thin Layer Drying Using Machine Vision and Artificial Neural Networks.
- [16] Sakib, S., Ashrafi, Z., Siddique, M. and Bakr, A. (2019) Implementation of Fruits Recognition Classifier Using Convolutional Neural Network Algorithm for Observation of Accuracies for Various Hidden Layers.
- [17] Nosseir, A. and Ahmed, S.E.A. (2019) Automatic Classification for Fruits' Types and Identification of Rotten Ones Using k-NN and SVM. *International Journal of Online and Biomedical Engineering*, **15**, 47. <https://doi.org/10.3991/ijoe.v15i03.9832>
- [18] Ummature, S.B. and Hanchinal, S.M. (2020) Multi Features Based Fruit Classification Using Different Classifiers. <https://doi.org/10.51201/jusst12503>

- [19] Zeeshan, M., Prabhu, A., Arun, C. and Rani, N.S. (2020) Fruit Classification System Using Multiclass Support Vector Machine Classifier. 2020 *International Conference on Electronics and Sustainable Communication Systems*, Coimbatore, 2-4 July 2020, 289-294. <https://doi.org/10.1109/ICESC48915.2020.9155817>
- [20] Behera, S.K., Rath, A.K., Mahapatra, A. and Sethy, P.K. (2020) Identification, Classification and Grading of Fruits Using Machine Learning and Computer Intelligence: A Review. *Journal of Ambient Intelligence and Humanized Computing*, 1-11.
- [21] Jaramillo-Acevedo, C.A., Choque-Valderrama, W.E., Guerrero-Álvarez, G.E. and Meneses-Escobar, C.A. (2020) Hass Avocado Ripeness Classification by Mobile Devices Using Digital Image Processing and ANN Methods. *International Journal of Food Engineering*, **16**, 1-8. <https://doi.org/10.1515/ijfe-2019-0161>
- [22] Ogbaji, M. and Iorliam, I.B. (2020) Effect of Drumstick Tree (*Moringa oleifera*) and Neem (*Azadirachta indica*) Leaf Powders on Shelf Life and Physiological Quality of Okra Fruits during Storage. *International Journal for Research in Biology & Pharmacy*, **6**, 1-12. <https://gnpublication.org/index.php/bp/article/view/1214>
- [23] Kumah, P., Olympio, N.S. and Tayviah, C.S. (2011) Sensitivity of Three Tomato (*Lycopersicon esculentum*) Cultivars—Akoma, Pectomech and Power—To Chilling Injury. *Agriculture and Biology Journal of North America*, **2**, 799-805. <https://doi.org/10.5251/abjna.2011.2.5.799.805>
- [24] AOAC (2007) Association of Official Analytical Chemists. Official Methods of Analysis. 15th Edition, Washington DC, 49.
- [25] Ibitoye, A.A. (2005) Laboratory Manual on Basic Methods in Plant Analysis. Concept IT and Educational Consults, Akure, 2-3.