**Elevating Apple Grading: The Power of Advanced Algorithms in Orchard Management**

**Pankaj Shahare(**[**2021bit058@sggs.ac.in**](mailto:2021bit058@sggs.ac.in)**), Dr. A.D. Sawarkar(**[**adsawarkar@sggs.ac.in**](mailto:adsawarkar@sggs.ac.in)**)**

1. Department of Information Technology, Shri Guru Gobind Singhji Institute of Engineering and Technology Nanded
2. Department of Information Technology, Shri Guru Gobind Singhji Institute of Engineering, and Technology Nanded.

**Abstract:**

This paper introduces a technique to enhance the accuracy of predicting apple pleasant with the aid of thinking about different factors along with length, weight, acidity, ripeness, crunchiness, juiciness, sweetness, and greater. These characteristics are essential markers of the best apples and are crucial in determining whether an apple is deemed appropriate or not. To attain this, we rent gadget mastering algorithms like logistic regression, choice trees, okay-nearest neighbours, help vector machines, naive Bayes, and random forests. These algorithms analyse the connection among these attributes and the overall high-quality of the apple.

Furthermore, we enhance the accuracy of our predictions by means of incorporating move-validation and grid seek strategies. While grid search assists us in finding the ideal set of model parameters to increase prediction accuracy, cross-validation enables us to confirm how well our models generalize to new data.

After thorough experimentation, we determined that the random wooded area algorithm, when blended with pass-validation and grid search, outperformed other algorithms, and done the very best accuracy of 90%. This incorporated technique provides a robust approach for appropriately assessing the pleasant of apples, that's precious for farmers and stakeholders inside the agricultural and food industries.

**Keyword:** Apple quality prediction, machine learning algorithms, cross-validation, grid search, random forest, accuracy enhancement, agricultural industry**.**

1. **Introduction:**Apples have gained global popularity, especially in nations that produce them in great amounts like the United States, China, and Poland. People are requesting more apples because they are becoming more aware of the fruit's health benefits recently [1]. In addition, new technologies in storage have been developed that allow apples to be stored at low temperatures for several months without losing their freshness. one example is 1-MCP (1-methylcyclopropene) [2]. Thanks to advancements in preservation techniques, apples can now be enjoyed throughout the year instead of only during their typical growth seasons. However, despite these improvements, the apple industry still has its struggles. Climate change is affecting how apples grow, and there's a need for better methods to control production. When there's too much production during times when apples aren't typically grown, it can lead to a drop in quality, making it harder for producers to maintain consistent standards[3].

To maintain their profitability and expedite the apple quality prediction process, farmers need to implement several strategies. Firstly, they can diversify their apple varieties to appeal to different markets and consumer preferences. Secondly, they should focus on efficiently packing apples according to their quality, reducing the time needed for sorting and packing. Lastly, they can explore ways to minimize labour costs by automating the separation of high-quality and low-quality apples, ensuring that only the best produce reaches the market. In the cutthroat apple market, these tactics help preserve profits while simultaneously increasing efficiency [4].

We've developed a cutting-edge apple quality prediction model that not only saves time in sorting apples but also reduces labour costs. This model considers essential attributes like size, sweetness, crunchiness, juiciness, ripeness, and acidity to accurately assess apple quality. By inputting these attributes, farmers can quickly determine the quality of each apple, streamlining the sorting process and minimizing the need for manual labour.

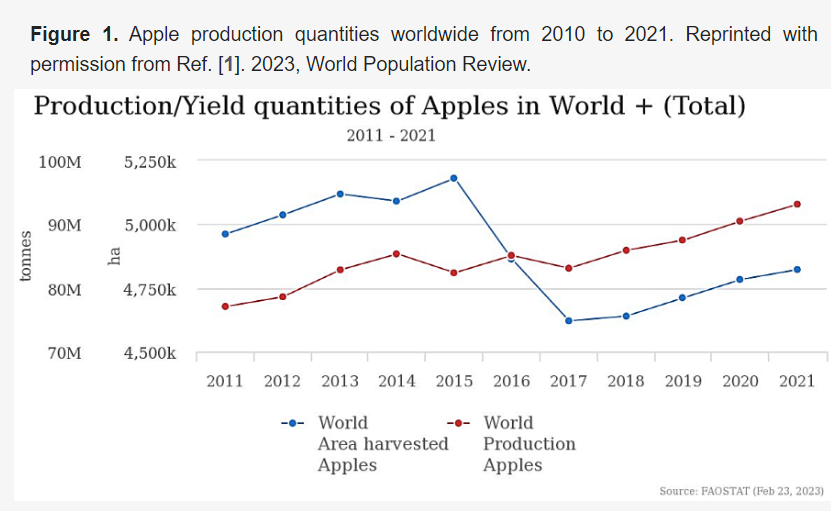
Moreover, our model isn't just for farmers. It's also handy for consumers who want to check the quality of random apples they purchase from the market. Simply input the attributes of the apple, and our model will provide an instant quality assessment, ensuring you always get the best produce. With our innovative approach, we're revolutionizing apple quality assessment, making it faster, more efficient, and cost-effective for everyone involved**.**

1. **Literature Review:**

In the past, predicting apple quality relied on manual inspection by farmers, which was subjective and prone to errors. Machine learning methods now analyse large datasets of apple attributes, offering more objective results. These methods automate the process, making it faster and scalable, leading to more reliable outcomes. This shift from manual to automated prediction has revolutionized apple quality assessment, ensuring consistent quality in today's competitive market [5].

Before the advent of advanced technology and machine learning algorithms, predicting apple quality relied heavily on manual inspection and the expertise of farmers and inspectors. The following are some conventional techniques that the older generation employed to forecast apple quality:

Traditionally, apple quality assessment relied on a combination of methods [6]. Visual inspection involved scrutinizing apples for ripeness, colour, size, and defects, demanding skilled personnel. Subsequent manual sorting categorized apples based on appearance, a labour-intensive process. Taste testing, though subjective, gauged flavour and overall quality, while basic measurements like weight and firmness provided additional insight. Moreover, seasoned professionals drew on their experience and intuition to make informed judgments Even though they were subjective, these techniques combined to provide a thorough assessment of apple quality by combining human judgment with objective measures.

****

Machine learning methods utilize algorithms to analyse extensive datasets containing diverse attributes of apples. Through this analysis, these methods discern patterns and relationships within the data, enabling them to generate predictions regarding apple quality. What distinguishes machine learning methods is their objectivity and reliance on data-driven insights, [7] which contribute to heightened accuracy and consistency in predictions. These techniques continuously improve their comprehension and adjust to new information by learning from vast collections of labelled data, which eventually increases their predictive power. This data-driven approach not only bolsters the reliability of predictions but also facilitates more informed decision-making in apple quality assessment, ultimately leading to improved outcomes in orchard management and fruit distribution.

Research in predicting apple quality using categorization methods from machine learning is still limited [7]. Many studies have not considered all the pertinent variables that can impact the quality of apples, such as the surrounding environment. Storage methods, and treatments after harvesting. Standardized techniques for gathering data, getting it ready for analysis, and assessing prediction models are also lacking. It's still difficult to comprehend these models' inner workings and to provide an explanation for their choices, particularly in the case of more complicated ones like ensemble approaches and deep learning. Additionally, there's been little exploration into creating models that can adapt to changes in the apple production and supply chain in real-time. These problems show how more thorough study and established procedures are required to forecast apple quality using classification approaches [8].

**Exiting Research Paper:**

Existing research on machine learning classification has explored various algorithms and methodologies.

**Paper-1:** Identification of apple varieties using hybrid transfer learning and multi‑level feature extraction Serhat Kılıçarslan1 Emrah Dönmez1  Sabire Kılıçarslan[9].

Model Summary : Accurate identification of apple varieties is crucial in pomology and agriculture, impacting orchard management, consumer satisfaction, and the economic viability of apple cultivation. Traditional identification methods are prone to errors, necessitating advanced technologies like image processing and machine learning. This study employs feature extraction techniques, including MobileNetV2, EfficientNetV2B0, and a combination of GLCM and Color-Space algorithms, followed by machine learning models (SVM, KNN, RSS, and Random Forest) to classify apple varieties. The best results were achieved using "EfficientNetV2B0+GLCM+Color-Space" with ReliefF feature selection and the Random Forest algorithm, achieving 98.33% accuracy. The study demonstrates that combining deep learning and traditional features significantly enhances the accuracy and efficiency of apple variety classification, supporting better orchard management and sustainable agricultural practices.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Algorithm | Accuracy(in %) | precision(in %) | Recall(in %) | F1-Score(in %) | MAE |
| SVM | 95.50 | 95.45 | 95.15 | 94.95 | 0.2420 |
| KNN | 91.2 | 91.40 | 91.10 | 91.15 | 0.0509 |
| RSS | 87.40 | 87.30 | 87.10 | 86.90 | 0.1076 |
| RF | 92.20 | 92.10 | 92.00 | 92.04 | 0.1041 |

**Paper-2:**Classification of Apples using Machine Learning by Agus pratonda Devira M.A.Harshap[10].

Model Summary : The study focuses on developing an automatic apple classification model using k-nearest neighbors (KNN) and support vector machine (SVM) algorithms. Utilizing images of apple variants (Envy, Fuji, Malang, and Gala), which are converted to grayscale and resized for efficiency, the model aims to improve reliability over manual classification. The results show high accuracy rates of 94.00% for KNN and 94.50% for SVM, indicating promising potential for applications in the agricultural processing industry. This approach addresses the limitations of human senses in large-scale apple classification and supports the efficient processing and distribution of apples.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fold | K=1 | K=3 | K=5 | K=7 | K=9 |
| 1 | 100 | 95 | 95 | 95 | 100 |
| 2 | 90 | 85 | 85 | 90 | 85 |
| 3 | 95 | 85 | 90 | 90 | 80 |
| 4 | 100 | 95 | 90 | 95 | 95 |

Fig. 2 show summary of experiment using KNN.

|  |  |  |  |
| --- | --- | --- | --- |
| fold | linear | poly | rbf |
| 1 | 95 | 90 | 100 |
| 2 | 85 | 85 | 85 |
| 3 | 95 | 90 | 90 |
| 4 | 100 | 100 | 95 |

Fig. 3 show summary of experiment using SVM.

**Paper-3:** Apple Varieties Classification Using Deep Features and Machine Learning by Alper

Taner,Mahtem Mengstu,Kemal Selvi,Huseyin Duran[11].

Model Summary: This study explores the automatic classification of apple varieties using computer vision and machine learning techniques to address the limitations of manual sorting. Initially, transfer learning with seven CNN architectures (VGG16, VGG19, InceptionV3, MobileNet, Xception, ResNet150V2, DenseNet201) was employed, with DenseNet201 achieving the highest accuracy (97.48%). Further, deep features extracted using DenseNet201 were used to train traditional ML models (SVM, MLP, RFC, KNN), with SVM achieving the best performance at 98.28%. Additionally, the study investigated the impact of principal component analysis (PCA) on classification performance, finding that MLP achieved the highest accuracy (99.77%). These findings demonstrate the effectiveness of combining deep learning and traditional machine learning in enhancing apple variety classification, supporting the agricultural industry's need for efficient and accurate sorting systems.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Accuracy | Precision | Recall | F1-score |
| SVM | 98.28 | 98.32 | 98.10 | 98.18 |
| RFC | 91.86 | 91.65 | 91.69 | 99.56 |
| MLP | 98.05 | 97.99 | 96.47 | 99.96 |
| KNN | 89.33 | 91.59 | 98.46 | 98.59 |

Fig. 4 show summary of experiment

1. **Methodology:**

The dataset, sourced from Kaggle, comprises various attributes of apples commonly used in quality assessment. These attributes include size, weight, acidity, ripeness, crunchiness, juiciness, and sweetness. Each attribute is measured or assessed on a numerical scale, providing quantitative insights into the characteristics of the apples. The dataset is organized to make it easier to use machine learning algorithms for analysis and apple quality prediction [12].

A correlation matrix is a symmetric matrix that displays the correlation coefficients between pairs of variables in a dataset. The correlation coefficient, denoted by "r," measures the strength and direction of the linear relationship between two variables [13].

The formula to calculate the correlation coefficient (r) between two variables, X and Y, is given by:

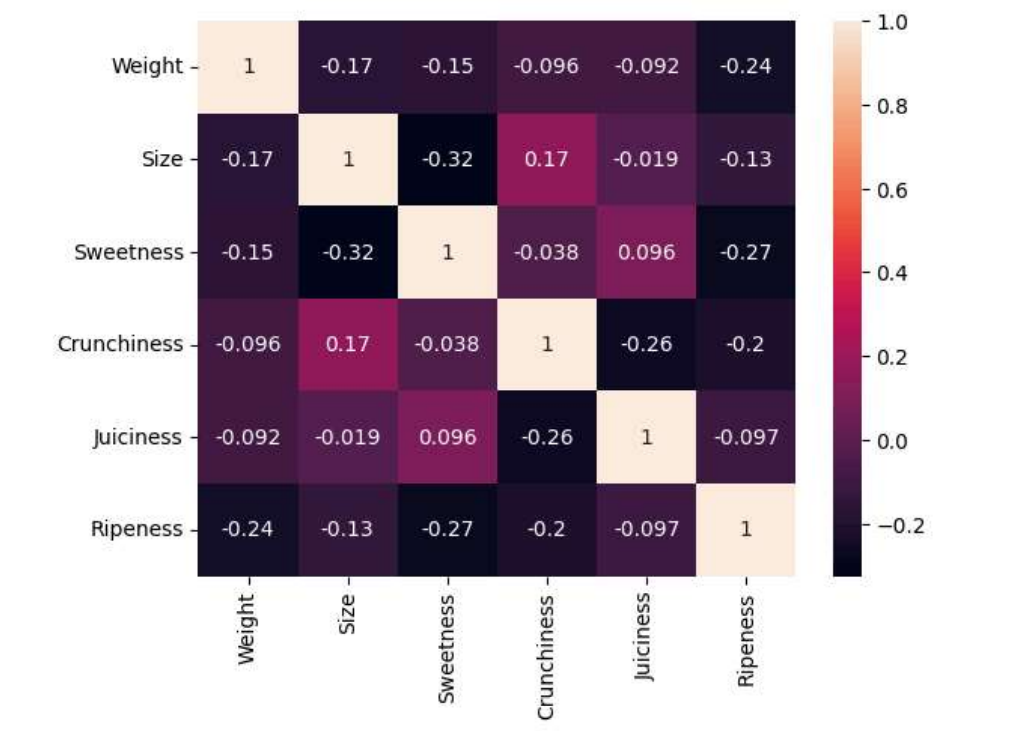


Fig No.5 correlation matrix (shows relationship between each attribute)

Similar to slicing apples for a delicious pie, data cleansing is essential for machine learning applications. We wouldn't utilize leaves or rotting apples, and our data is no different! During the first step, we carefully review the data, searching for mistakes such as missing information, illogical values, and typos. Then, if the problems are too serious, we remove the data points completely. If not, we repair the faults and use suitable techniques to fill in the missing portions. This guarantees the completeness and accuracy of all the data we utilize [14].

The data must then be shaped such that the machine learning model can exploit it to its fullest potential. Similar sized apple slices are an example of how normalization guarantees that all features are on the same scale. By using feature engineering, we may produce even more useful features, such as a sweetness-acidity ratio, which is far more useful for forecasting apple quality than individual sweetness and acidity values. Additionally, we can focus on the features that are most important to achieving our goal by eliminating less important features from the data. We make sure our machine learning model has the best quality data to work with by adhering to these data cleaning practices, which produce the most accurate and dependable results [15].

By using these methods for cleaning data, we make sure our machine learning model gets the best apples (data) to work with, which produces predictions that are more accurate and dependable.

**3.3.1 Machine Learning Model:**

1. **Logistic Regression:** Selected for its simplicity, interpretability, and effectiveness in binary classification tasks. Suitable for datasets with linearly separable features [16]

*Where,*

*z*=*β*0​+*β*1​*x*1​+*β*2​*x*2​+⋯+*βn*​*xn*​

1. **Decision Trees:** The decision tree algorithm is a non-parametric supervised learning method used for classification and regression. The process of constructing a decision tree involves recursively partitioning the feature space into homogeneous subspaces based on the input features. The decision tree uses a tree structure to model the decisions and their possible consequences. [17].

The core part of a decision tree is how it decides to split the nodes. Common criteria include Gini impurity, entropy (information gain), and others:

Gini Index =

Entropy =

Information Gain = Entropy(D) -

Where 𝑝𝑖 is the proportion of instances belonging to class *i* in the dataset *D*, and 𝐶is the number of classes.

1. **Random Forest:** Utilized for its ensemble nature, which combines multiple decision trees to improve predictive performance and robustness against overfitting. Offers feature importance analysis and can handle large datasets efficiently [18][19].
2. **Support Vector Machine (SVM):** Support Vector Machine (SVM) is a powerful supervised learning algorithm used for classification and regression tasks. In classification, SVM aims to find the optimal hyperplane that separates the data points of different classes in the feature space with the maximum margin. The decision boundary is determined by support vectors, which are the data points closest to the hyperplane. [20].

(*x*)=*β*0​+*βTx*

Where:

* 𝑓(𝑥)*f*(*x*) is the decision function.
* 𝑥*x* is the input vector (feature vector).
* 𝛽*β* is the weight vector.
* 𝛽0*β*0​ is the bias term.

The predicted class for a given input 𝑥*x* is determined by the sign of 𝑓(𝑥):

If f(x)>0 predict class 1,

If f(x)<0 predict class 0

1. **K-Nearest Neighbours (KNN):** The k-Nearest Neighbors (k-NN) algorithm is a simple and effective supervised learning algorithm used for classification and regression tasks. In k-NN classification, the class label of a data point is determined based on the majority class among its k nearest neighbors in the feature space [21].

**Euclidean Distance:**

2 .

**k-NN Algorithm Steps**:

1. Calculate distances between 𝑥new and all training data points.
2. Select the k nearest neighbors of 𝑥new.
3. Determine the class label of xnew based on majority voting among its k nearest neighbors.

**3.3.2 Model Workflow :**

The workflow of the model training and evaluation process as depicted in the flowchart involves several key steps:

A diagram of a software development process

Description automatically generated

Fig. 6 workflow of the model

A machine learning process for identifying different apple varieties. This process involves two crucial stages: feature extraction and feature selection. Let's delve deeper into each stage and understand how they contribute to accurate apple classification.

**Feature Extraction: Transforming Raw Data into Machine-Readable Format**

Imagine you have a basket full of apples of various shapes, sizes, and colours. To identify the varieties, your eyes and brain naturally extract features like colour, shape, texture, and size. Similarly, in machine learning, feature extraction plays a vital role in transforming raw data (images of apples in this case) into a format that a machine learning model can understand and analyse. The specific techniques used for feature extraction depend on the type of data and the desired outcome. In the case of apple variety identification, common feature extraction methods include[22]:

* **Colour analysis:** Extracting colour features like average colour, dominant colour, and colour distribution helps distinguish between varieties with different colour profiles.
* **Shape analysis:** Analysing the shape of the apple using techniques like geometric moments or contour analysis allows the model to differentiate between round, elongated, or irregular shapes.
* **Texture analysis:** Extracting texture features like smoothness, roughness, and presence of blemishes helps identify varieties with distinct textures.
* **Size analysis:** Measuring the size of the apple using techniques like area or diameter can be helpful in distinguishing between larger and smaller varieties.

These extracted features become the building blocks for the next stage: feature selection.

**Feature Selection: Choosing the Most Relevant Features**

Not all extracted features are equally important for accurate classification. Some features might be redundant or irrelevant, potentially leading to overfitting and reduced performance of the model. Therefore, the next crucial step is featuring selection.Feature selection involves choosing the most informative and relevant features from the pool of extracted features. This is often done using statistical methods like correlation analysis, information gain, or chi-square tests. These methods help identify features that have the strongest correlation with the target variable (apple variety in this case) and eliminate those that contribute little or negatively to the model's performance.By focusing on the most relevant features, the model can learn more effectively and make more accurate predictions[22].

**Machine Learning Algorithms: Making Predictions Based on Selected Features**

Once the relevant features are selected, they are fed into various machine learning algorithms to train a model for apple variety identification. The diagram typically shows different algorithms like Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Decision Trees, and Random Forests. These algorithms learn from the training data and build models that can predict the variety of a new apple based on its extracted and selected features.

**Performance Evaluation: Measuring the Model's Accuracy**

The final step involves evaluating the performance of the trained model. This is typically done using metrics like accuracy, precision, recall, and F1-score. These metrics help assess how well the model can correctly identify different apple varieties.By iteratively refining the feature extraction and selection processes, and fine-tuning the chosen machine learning algorithms, the accuracy of the model can be improved. This allows for more reliable and efficient identification of apple varieties, which can benefit various stakeholders in the agricultural industry[23].



**3.3.3 Evaluation matrix:**

By contrasting the models' predictions with the dataset's actual class labels, evaluation metrics are used to evaluate the effectiveness of classification models. These are a few popular metrics for assessing categorization models:

1. **Accuracy:** Accuracy measures the proportion of correctly classified instances out of the total instances. It is calculated as the ratio of the number of correct predictions to the total number of predictions.

*Accuracy*=(TN+TP)/(TN+TP+FN+FP)

1. **Precision:** Precision measures the proportion of true positive predictions among all positive predictions made by the model. It focuses on the correctness of positive predictions.

Precision =TP/(TP+FP)

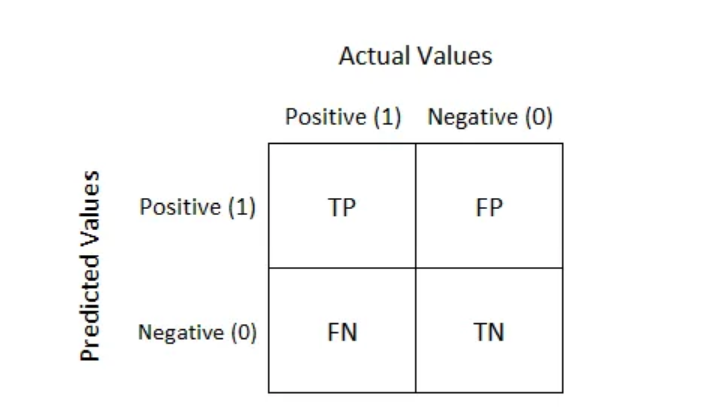
1. **Recall (Sensitivity):** Recall measures the proportion of true positive predictions among all actual positive instances in the dataset. It focuses on the model's ability to correctly identify positive instances.

Recall =TP/(TP+FN)

1. **F1-Score:** F1-score is the harmonic mean of precision and recall. It provides a balanced measure of the model's performance, taking both false positives and false negatives into account.

F1-Score= 2\*((precision \* recall)/ (precision + recall))

1. **Confusion Matrix:** A confusion matrix is a table that summarizes the model's performance by comparing predicted class labels with actual class labels. It provides insights into the true positives, true negatives, false positives, and false negatives [24].



1. **Experimental Setup:**

We used the scikit-learn Python package to create a variety of machine learning models for classification tasks. The following models of classification were applied:

1. Logistic Regression
2. Support Vector Machine (SVM)
3. Random Forest
4. Decision Tree
5. K-Nearest Neighbours (KNN)
6. Naive Bayes

These models were chosen because to their ability to handle classification tasks well and their adaptability to various dataset types. Python was selected as the programming language because of its wide compatibility with machine learning frameworks and its user-friendliness. Computational resources utilized included standard personal computing hardware.

In data analysis, pandas (pd) are a go-to tool for organizing and cleaning up data, especially when dealing with tables. train\_test\_split is like splitting your study notes into parts for learning and testing. Label Encoder translates words into numbers, making it easier for computers to understand [25]. Classification algorithms, such as Random Forest Classifier and Logistic Regression, sort data into groups. Evaluation measures, such as the categorization report and accuracy score, aid in gauging the effectiveness of these algorithms.

To improve the performance of the classification models for apple quality prediction, hyperparameter tuning was done. A range of hyperparameter values was methodically investigated for every model using a grid search technique. By using cross-validation, the grid search sought to find the best set of hyperparameters to maximize performance measures including F1-score, accuracy, precision, and recall. The models were optimized through this iterative method to yield optimal performance on the classification test.

To ensure robustness and generalizability in predicting apple quality, three-fold cross-validation was employed for the classification models. Three folds, or subsets, of the dataset were used to train the model three times. Each iteration involved training on two folds and validating on the remaining fold. To get an overall estimate of the model's performance, this process was repeated for each fold, and performance measures were averaged over all folds. Using three-fold cross-validation strikes a balance between computational efficiency and robustness in evaluating the classification models [26].

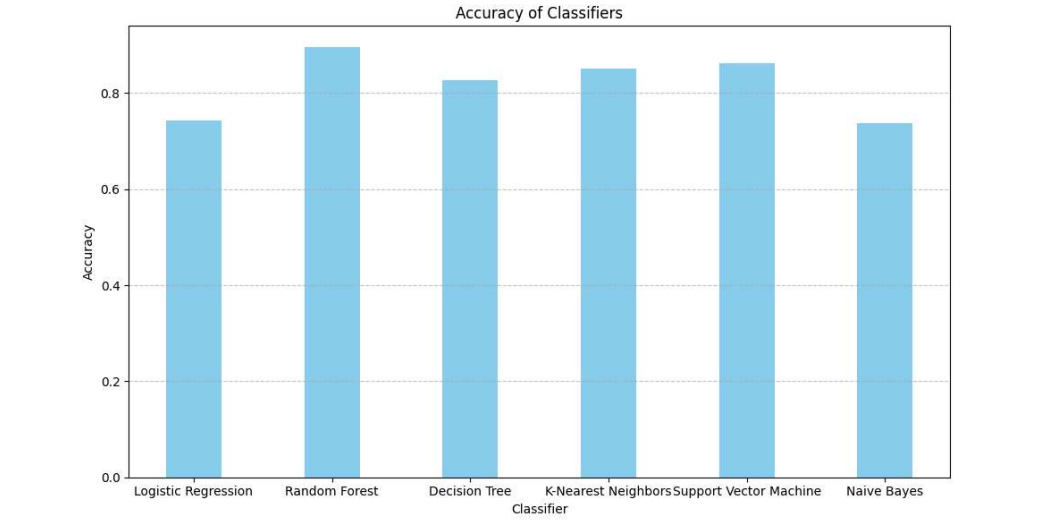


Fig.7 comparing the accuracy of all algorithms.

1. **Results :**

The true values from the dataset are contrasted with the predicted values generated by our classification model in Table 1's prediction error plot. This visual aid allows us to see how much the predicted values differ from the actual ones, giving us a clearer understanding of how well the model performs. By examining these differences, we can evaluate the precision and potency of our model in predicting the qualities of apples.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Algo Name | Accuracy (in %) | Precision | Recall | F1-Score |
| Random Forest | 90.5 | 0.90 | 0.91 | 0.91 |
| Logistic Regression | 74.375 | 0.75 | 0.73 | 0.74 |
| Decision Tree | 82.625 | 0.83 | 0.82 | 0.83 |
| K-Nearest Nei. | 82.625 | 0.83 | 0.85 | 0.85 |
| SVM | 86.25 | 0.86 | 0.86 | 0.86 |
| Naïve Bayes | 73.75 | 0.74 | 0.74 | 0.74 |

Fig 8. shows the summary of the model evalution

**A graph with green bars

Description automatically generated**

Fig. 9 show the comparison of precision

**A graph of a graph

Description automatically generated with medium confidence**

Fig.10 show the comparison of Recall

**A graph of a graph

Description automatically generated with medium confidence**

Fig. 11show the comparison of F1-score

1. **Discussion:**

When it comes to predicting the quality of apples, one machine learningmodel reigns supreme: the Random Forest. It's like the superstar of apple prediction algorithms! Let me break down why it's so awesome in simple terms:

Random Forest is like a superhero among apple sorters. It boasts unmatched accuracy, rarely making mistakes when determining apple quality, with an impressive 90.5% accuracy rate. Unlike other models, such as Logistic Regression and Decision Trees, Random Forest finds the perfect balance between precision and recall, ensuring it identifies both top-quality apples and those that aren't up to par without confusion. This is thanks to its ensemble approach, acting as a team of super apple sorters, combining the insights of many decision trees to make the final call. Through grid search and cross-validation techniques, Random Forest ensures robustness and generalizability, finding the best settings for decision-making and handling new scenarios effectively.

In simple terms, Random Forest is like having a bunch of super smart apple experts working together to tell you which apples are the best. Its amazing accuracy, ability to balance details, teamwork approach, and careful techniques make it a top choice for farmers and anyone else who wants to know their apples' quality.

**Conclusion:**

This study introduces a method to predict the quality of apples by considering various factors like size, weight, acidity, ripeness, crunchiness, juiciness, and sweetness. To examine the relationship between these variables and apple quality, they employed a variety of machine learning algorithms, including logistic regression, decision trees, k-nearest neighbours, support vector machines, naive Bayes, and random forests. To make their predictions more accurate, they used techniques like cross-validation and grid search, which help explore different combinations of parameters and assess how well the models perform on new data.

Out of all the algorithms they tested, the random forest algorithm performed the best, with an accuracy of 90.1%. They further improved the accuracy to 90.5% by using a technique called three-fold cross-validation within grid search. This means they divided the data into three parts and tested the model three times, rotating which part was used as the test set each time. This showed that their model was not only accurate but also robust and could generalize well to new data.

Overall, this integrated approach provides a reliable method for farmers and stakeholders in the agricultural and food industries to accurately assess the quality of apples. This could be incredibly useful for making decisions about harvesting, storage, and distribution, ultimately benefiting both producers and consumers.

**Reference:**

1. [Wikipedia](https://en.wikipedia.org/wiki/Apple)
2. **Fruit Ripening: Physiology, Signalling and Genomics** (Edited by Nath, M., Bouzayen, M., Mattoo, A. K., & Pech, J. C.)
3. Climate Change and Fruit Production (Edited by Fouad M. Hamada and Amadou M. Ballo)
4. [A robot system for the autodetection and classification of apple internal quality attributes](https://doi.org/10.1016/j.postharvbio.2021.111615)
5. [Classification Applications with Deep Learning and Machine Learning Technologies](https://d.docs.live.net/0ec779e978f91a47/Desktop/ML/Classification%20Applications%20with%20Deep%20Learning%20and%20Machine%20Learning%20Technologies)
6. [Mastering Classification Algorithms for Machine Learning: Learn how to apply Classification algorithms for effective Machine Learning solutions Partha Majumdar](https://www.google.com/search?sca_esv=de93c5c4281f23a7&sca_upv=1&cs=1&q=Mastering+Classification+Algorithms+for+Machine+Learning:+Learn+how+to+apply+Classification+algorithms+for+effective+Machine+Learning+solutions++Partha+Majumdar&stick=H4sIAAAAAAAAAGVUv4_cRBReb5TVni-RbveEiK4yR7Ncsx7ba--G4ohCJIqcQOEKKox_jX-NPR7PxN5xgVD-gihSKNIQUYHEH4AoUIREQQqKKxASEgUUiAYKmogKL1mPN-DKn9_7vve9N543vnx8dR7OAaDVsqgA97SZi3FKFQwVDzmUxjD2HBbjXMkcL4rzQEGBU-ZxHj6VnjPdmhPGXb_DzPNoiVzjqXRlg6NQteyF3nThrpDAa7NlL5cdjlQ9hZVKtnTNzczIZkkX9iwDIjvLRXWLANvKWZduJEZGQ6s3V1CSh2gn3XTrpKXL82yuGral1YsuiEhaJhx2Wrlu1rqpusJJDkO3irpsSJqcWEkhsJpWtlEJpyw1C7KqdNHYqmKA5Nq2tOvaxNdE0EuhTgnfBs3UWpS8H4FW1UthMzL9fL1I695Xa4SDHhaqHzKRDVYgy4xmq5wv_ARou-PxysTscAhbW1FKxUQgMUhVmJ22BtwoC0UYJpSjZa0KOUZ4oaP1jtNC5404rMRgbtQw0uGUlKoNPCDK2-um4rQS-qBkOnPtDvtr4q0tqPWteojR_rA5JBxBvtNdCXO7H5zOIh-sejt-2oRVtCOnwSISdA_UIYloP0loEn25ED96yJfGwqaloJdRgcLmF-mZtH_w-7OfD4_-kB588d0P0m-SfHAbYxogfidADgv8czx9WR7dylnM-PTq0b68tzmbhdcQML0l778bsHN8hv0Y8qk5NeS9syBzg5K-DaevyvJNjFDgbe7k9KWjQ3ky98SH-b-X9_pwNjwm2vtfXXz8_ei9yaB9Zq_deeNodjKRR2_izInzyVsXf39y8defpyeH8vjcWeMcZ3zy6Ncv49c__Pz0-JW9lvN4-OjH0w37o5--XT25dkkZzgaghR9oD0-enAzuD6V7X3_2zWg8lg4G2nA8aAZXPr38WDpzKAvKdkEoN19cIDdQiMuYRRlVIC6Vs-0-ub3dJ9efvykRrhWGFacoEP-vhPOiRADhpu8q-J-YQjG6u6FQRXnHKVnktCnJ3cx3yvsj6R81SAmg9gQAAA&sa=X&ved=2ahUKEwi-jczU0e-FAxVW1DgGHQ31D5EQ7fAIegUIABC9BQ)
7. [The Hundred-page Machine Learning Book](https://www.google.com/search?sca_esv=de93c5c4281f23a7&sca_upv=1&cs=1&q=The+Hundred-page+Machine+Learning+Book&stick=H4sIAAAAAAAAAE1Uv2_TQBSOgxqlbis1qRAoUygSCpVQfLZjJzAUAZU6tEIqHZgw8Y_z77PPPpycB4T4CyokGLqAmEDiD0AMqEJioANDB4SExAADYoGBBTGRQnyOt8_f-9773nt3V59bXeraXQAcQfJhJmCxo0eRn7Yj2DaCYZq60DWGxI1QOxwajousdmANE-Qi-5D7r9RHFBOqmwUmhpEmgS4fcovH2LEFVetJeUGnWT_OADUYHisTdb9f4MLIVC7qoeJoxCtoQ5VhoIWIVVcx0FREinDZk8PUVktzcYqRHcyEK_rIm8j5btgVZE0VR72CDLCfeBQWuZCkjCRF0JkTBG09c4poiHOEVS9mWPAzTc6YU-IrMR5kEmtskBGAkTgtresaNkVGGj6UUkynpOKrvYSWIxCzUZ_ZdBQTjXv-qPQ1MUJBCWPBtAmLBgMQhnI-zYx6pgfE2fEYiacU2IYTW46fsolALOMsVorcItCd0GY09FIa9EcCS0cwjaVgPOM0lmjOluXJRHdyggvs40TQgAFYeW2cZzTNWH6QEInoWoHNMTbGKhTLVo2ApOWyKcQ0gHSmuwQirRycRBwTDEo7pp_bmTOTToSxw-QGGNnYSctJQgVL_R476Dbtyz0tTZg8ceLAzr9wv7mF5e-_P6-0fnAPXrz7wH3j-OWtKEqtgO5YwZBY5m7UPMXXNhBxCW0utRb4-ePd9Iwcg-YGv3DDIrvRdmS6kDaVpszPb1uhbiXpddg8y_NXoyCwjOM72TzZWuEbXYP96P67vBerneoqFm-9Onr0vnazUZl8nfM7l1udtQZfuxaFQxc1No_-PD769XN9bYWv7w7HEYpC2tj_-tK9dPf5-uqZ-YnmSXX_4_qx-t6nt4OD0yfa1U4FTOBt8eHawVplr8rdf_3sTa1e55YrYrVeySuLT-fO7TpWe_MOMhPLvBAPbau9PX03tqbvRvvKxONejfsLmUT_cXsEAAA&sa=X&ved=2ahUKEwi0-Nfn0u-FAxXn4zgGHc8RA4kQ7fAIegUIABDdBQ)
8. [Apple: Cultivation, Diseases and their Management](https://www.ibpbooks.com/apple-cultivation-diseases-and-their-management/p/55514)
9. Identification of apple varieties using hybrid transfer learning and multi‑level feature extraction Serhat Kılıçarslan1 Emrah Dönmez1  Sabire Kılıçarslan
10. Classification of Apples using Machine Learning by Agus pratonda Devira M.A.Harshap
11. Apple Varieties Classification Using Deep Features and Machine Learning by Alper Taner,Mahtem Mengstu,Kemal Selvi,Huseyin Duran
12. [Kaggle dataset](https://www.kaggle.com/code/anzarwani2/apple-quality-classifier)
13. Understanding Correlation Matrices (Quantitative Applications in the Social Sciences) by Alexandria Hadd and Joseph Lee Rodgers
14. Feature Engineering for Machine Learning and Data Science by Enrico Santoro
15. **Logistic Regression** by David W. Hosmer Jr. and Stanley Lemeshow
16. Decision Tree Learning for Machine Learningby Andreas Géron
17. <https://link.springer.com/chapter/10.1007/978-1-4842-7150-6_14>
18. <https://simple.wikipedia.org/wiki/Random_forest>
19. Support Vector Machines for Pattern Recognition by Nello Cristianini and John Shawe-Taylor
20. Data Mining: Concepts and Techniques by Jiawei Han, Micheline Kamber, and Jian Pei
21. The Elements of Statistical Learning by Trevor Hastie, Robert Tibshirani, and Jerome Friedman
22. classifier workflow: step-by-step guide for the comparison of machine learning classification techniques by Zanifa Omary
23. Hands-On Machine Learning with Scikit-Learn, Keras &TensorFlowby Aurélien Géron
24. Python Machine Learningby Sebastian Raschka and Vahid Roostamzadeh
25. Handbook of Plant Disease Identification and Management
26. Playing with Nature History and Politics of Environment in North-East India