MANGO QUALITY PREDICTION WITH MACHINE LEARNING

Deepali A Kashalikar

Date: 15/05/2024

Abstract:

This report explores the application of machine learning (ML) techniques in automating the process of identifying export-quality mangoes without human involvement. As the mango export industry continues to thrive, traditional methods of quality assessment often rely on human judgment, which can be subjective and time-consuming. By harnessing ML algorithms, this study aims to develop a system capable of analyzing various factors such as size, color, texture, and blemishes to accurately classify mangoes as export-quality or otherwise. By eliminating the need for human intervention, this automated approach promises to enhance efficiency, reduce errors, and streamline the mango export process, ultimately contributing to increased productivity and profitability in the industry.

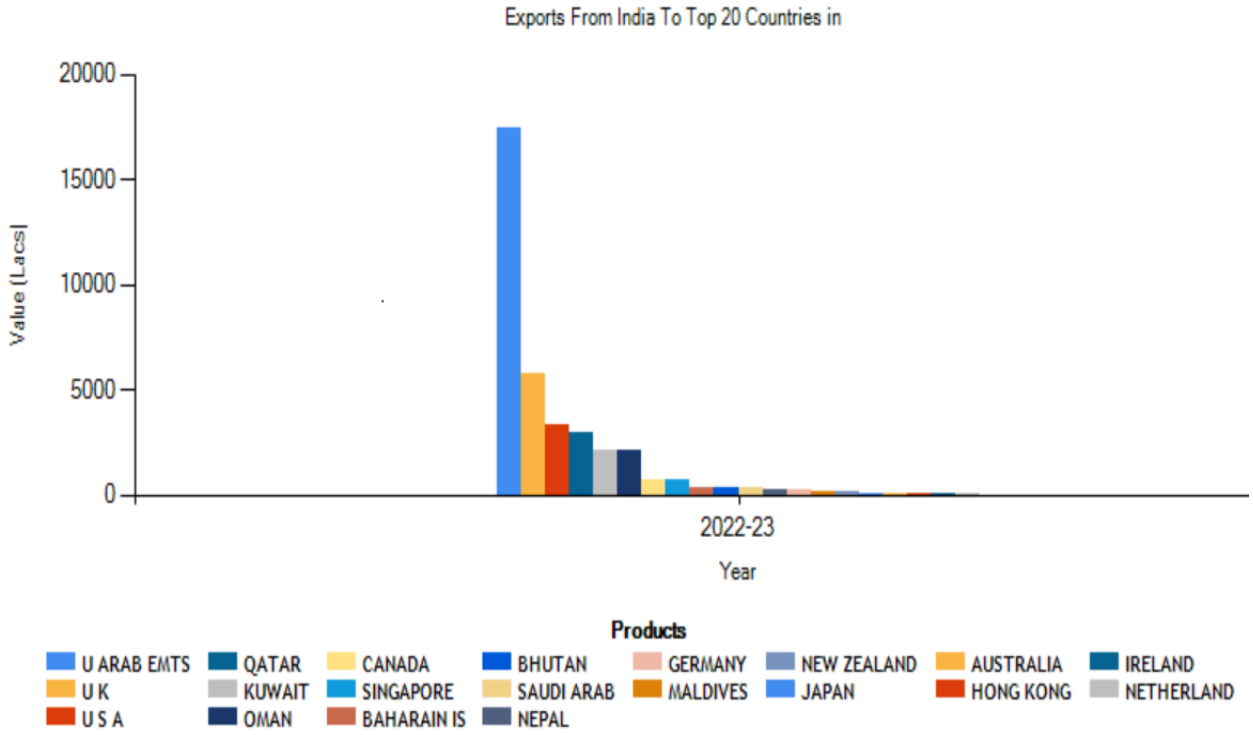
Problem Statement:

The mango export industry faces a challenge in ensuring consistent export-quality standards required to meet the growing demands of the industry, leading to potential losses in revenue. The quality of mangoes is influenced by various factors, including growth conditions, maturation before harvesting, and significantly, the post-harvest process. Traditional methods of manually grading and classifying mangoes are labor-intensive, leading to increased costs and inconsistent quality. Modern markets demand not only attention to external characteristics such as size, color, and firmness but also internal quality metrics like sugar content and acidity. While consumers often select fruits based on their appearance, their decision to repurchase depends on the fruit's taste. In light of these challenges, this study explores the application of Machine Learning (ML) to enhance the grading process of mangoes by analyzing both external features and internal quality indicators. This innovative approach aims to streamline the grading process, ensuring consistent quality and meeting the stringent requirements of both vendors and consumers.

Market/Customer/Business Need Assessment:

The market/customer/business need assessment involves evaluating the current landscape of the mango export industry and identifying the specific challenges and requirements that drive the need for a machine learning-based solution for mango quality assessment. This assessment encompasses several key aspects:

1. Market Demand and Trends: Understanding the current demand for mango exports and trends in the global market. This includes analyzing market growth projections, consumer preferences, and emerging market opportunities. the exports from India (2022-23)



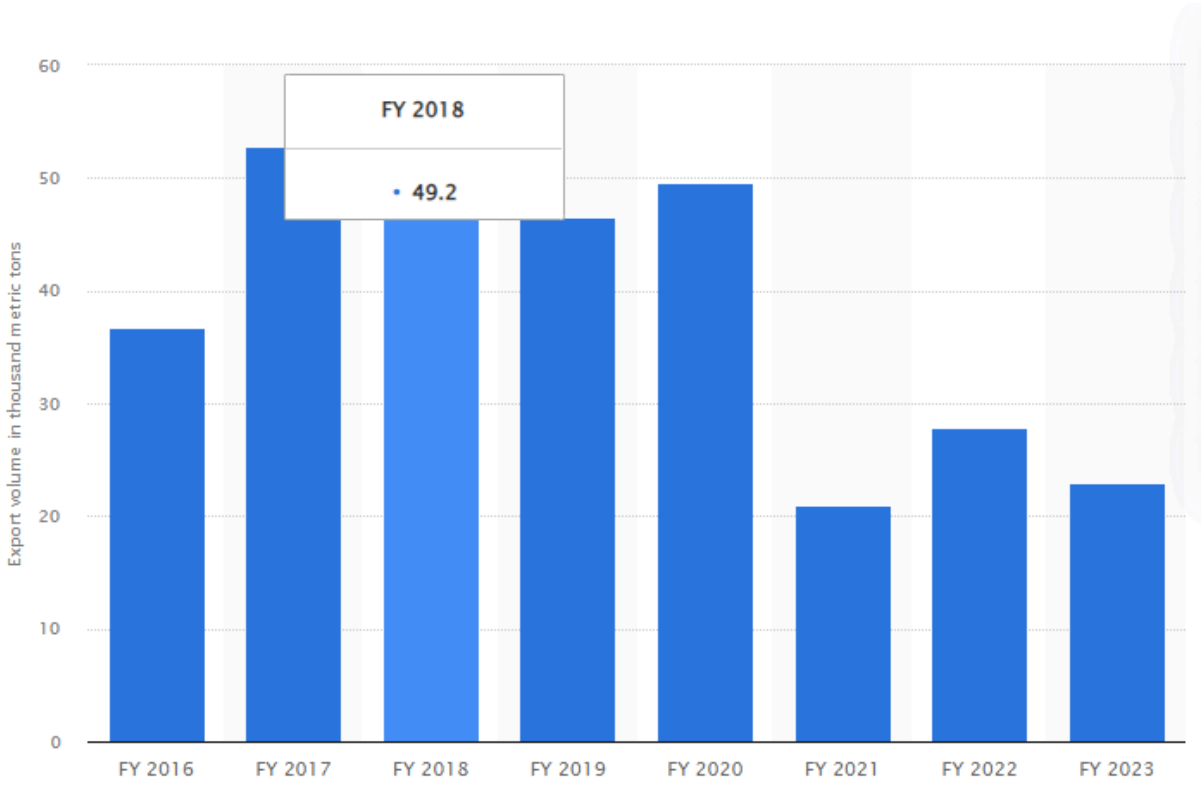
2. Competitive Analysis: Assessing the competitive landscape to identify the strengths and weaknesses of existing mango quality assessment methods employed by competitors. This involves evaluating the efficiency, accuracy, and scalability of competing solutions.

Table 1: Top 20 Mangoes, mangosteens and guava producing countries in 2019 and 2020

2019			2020		
Country Name	**Quantity (t)**	**% exported**	**Country Name**	**Quantity (t)**	**% exported**
India	25631000	0.57	India	24748000	0.52
Indonesia	3294817	0.04	Indonesia	3617271	0.64
China	2582791	0.34	China	2539717	0.42
China, mainland	2415000	0.00	Mexico	2373111	17.77
Mexico	2396675	17.21	China, mainland	2368180	0.00
Pakistan	2270229	4.28	Pakistan	2344647	4.57
Brazil	2002849	10.79	Brazil	2135304	11.40
Thailand	1643058	29.88	Malawi	1938066	0.01
Malawi	1492687	0.03	Thailand	1657589	23.61
Bangladesh	1456331	0.02	Bangladesh	1448396	0.02

3. Regulatory Compliance: Assessing the regulatory requirements and standards imposed by importing countries on mango exports. This includes understanding the documentation, certification, and quality control measures necessary to meet international trade regulations.

4. Business Objectives: Aligning the proposed machine learning solution with the business objectives of mango exporters. This involves considering factors such as cost-effectiveness, ease of integration into existing workflows, and the potential impact on productivity and profitability.



export volume in metric tonnes

5. Technological Feasibility: Evaluating the feasibility of implementing machine learning technology for mango quality assessment. This includes assessing the availability of data, the suitability of ML algorithms for the task, and the technical infrastructure required for deployment.

By conducting a comprehensive assessment of these factors, stakeholders can gain insights into the market/customer/business needs and make informed decisions regarding the development and implementation of a machine learning-based solution for mango quality assessment.

Target Specifications and Characterization:

A. Utilize predetermined datasets of mango quality parameters and export standards for predictive analysis: datasets serve as a foundation for our ML algorithms, enabling the development of robust predictive models. These models can then be used to analyze new batches of mangoes, ensuring they meet the necessary quality standards before

reaching the market. This predictive capability not only enhances the efficiency of the grading process but also ensures a consistent and high-quality product for consumers.

B. requirements not only on external factors like size, color, and firmness but on internal quality factors such as sugar content and ripeness
internal quality factors are equally crucial and include:
    a. Sugar Content: The sweetness of the mango, a primary indicator of taste.
    b. Acidity: The balance between sweetness and tartness, affecting the overall flavor profile.
    c. Ripeness: The stage of maturation, which influences both taste and texture.

    By incorporating both sets of factors into our ML models, we can provide a comprehensive assessment of each mango's quality.

C. Sorting mangoes in terms of maturity.: An essential aspect of mango grading is sorting the fruit based on its maturity. This process involves determining the stage of ripeness, which affects the mango's taste, texture, and shelf life. Using machine learning, we can analyze various indicators of maturity, such as color, firmness, and internal quality metrics, to accurately classify mangoes into different maturity categories.

External Search:

In understanding the need for implementing a Machine Learning (ML) system for mango quality detection and export optimization, several sources were consulted:

1. Understanding the parameters that define mango quality, such as color, size, texture, and ripeness, is crucial for ensuring export readiness and meeting international standards. US dept of Agriculture: Grades of Mangoes

2. Exploring how ML techniques have been applied in agriculture, particularly in crop quality assessment and yield optimization, provides insights into the potential benefits for the mango industry. https://indatalabs.com/blog/ml-in-agriculture

3. Export Optimization Strategies: Analyzing existing strategies employed by successful exporters to streamline the export process, enhance product quality, and meet market demands aids in designing an effective ML-based system for mango: machine learning in qualtity detection of fruits

Benchmarking:
Benchmarking against existing ML-based systems for fruit quality assessment and export optimization, such as those used in the citrus or apple industries, provides valuable insights into best practices, potential challenges, and opportunities for adaptation to the mango sector.

Human intelligence	**Machine intelligence**
Experience-based judgment	Dependence on data quality, Limited adaptability
experienced graders may achieve high accuracy but it's not guaranteed	Accuracy is typically high once properly trained and consistent
Slow speed due to manual process	Capable of processing a large lot quickly
Can assess size, color, and firmness through sensory perception	Uses image processing and computer vision to assess size, color, and firmness with high precision

Applicable Patents:
Exploring patents related to ML-based quality assessment systems for fruits, particularly mangoes, and bananas, can offer innovative ideas and technologies that may apply to the project. Patents focusing on image processing, spectroscopy, or sensor-based technologies are particularly relevant.

- MDPI journal on mango: https://www.mdpi.com/2076-3417/10/17/5775
- A Case Study On How Did Reliance Became The Largest Mango Exporter In India: link
- MDPI journal on Bananas: https://www.mdpi.com/2079-9292/11/24/4100

In the context of exporting mangoes, several government regulations and environmental considerations are pertinent to ensure compliance and sustainability:

Applicable regulations:

1. Certification and Documentation: Exporters may need to obtain certification or documentation to demonstrate compliance with regulatory requirements. This may include phytosanitary certificates, food safety certifications (e.g., HACCP), organic certifications, and documentation verifying compliance with trade regulations.

2. Sustainability Initiatives: Increasingly, consumers and importers are demanding sustainably sourced products. Mango exporters may opt to participate in sustainability initiatives, such as Fair Trade certification or organic farming practices

Do's	Don'ts
Registration of farmers by APEDA for export	Do not use unauthenticated planting materials.
Pre harvest spraying of recommended chemicals	Don't have any pesticide residue.
Maturity • Full mature – Export through air • 85-90% maturity – Export through sea.	Don't use defected, deformed, bruised and diseased fruits.
Mangoes shall comply with the residue levels of heavy metals, pesticides and other food safety parameters as laid down by the Codex Alimentarius Commission for exports.	Don't pack different varieties and sizes of fruits together.
Nearly 200 pesticides are now registered in India for use in agriculture under section 9 (3) of the Insecticides Act, 1968, which may come into our food chain from their intentional or unintentional use on the fruits as fruits are mostly consumed as fresh.	Do not use banned pesticides like: Aldrin, Aldicarb, Benzene hexachloride (BHC), Calcium cyanide, Chlordane, Chlorobenzilate, Dibromochloropropane, Copper acetoarsenite and Dieldrin
Vapour-heat treatment for export to Japan Temperature of the treatment chamber shall be raised step by step to 500C for 20 minutes.	Don't use water of inferior quality.
Irradiation at 400 Grays using Cobalt-60 for export to USA	Don't do it without the inspection of USA agencies.
Grade fruits according to weight or as per requirement of importing country	Don't under or over fill fruits during packaging.

Applicable constraints:

1. Data Collection and Standardization: Obtaining a comprehensive dataset of mango quality attributes from diverse sources is challenging. Standardizing data collection methods and ensuring data consistency across different regions and varieties is crucial for training a reliable machine learning model.

2. Technical Expertise: Developing and deploying a machine learning model requires expertise in data science, programming, and machine learning algorithms. Access to skilled data scientists and software engineers is essential for building and maintaining the predictive model and associated infrastructure.

3. Infrastructure and Resources: Implementing a machine learning solution needs adequate computing infrastructure and resources for data storage, processing, and model training.

Business model:

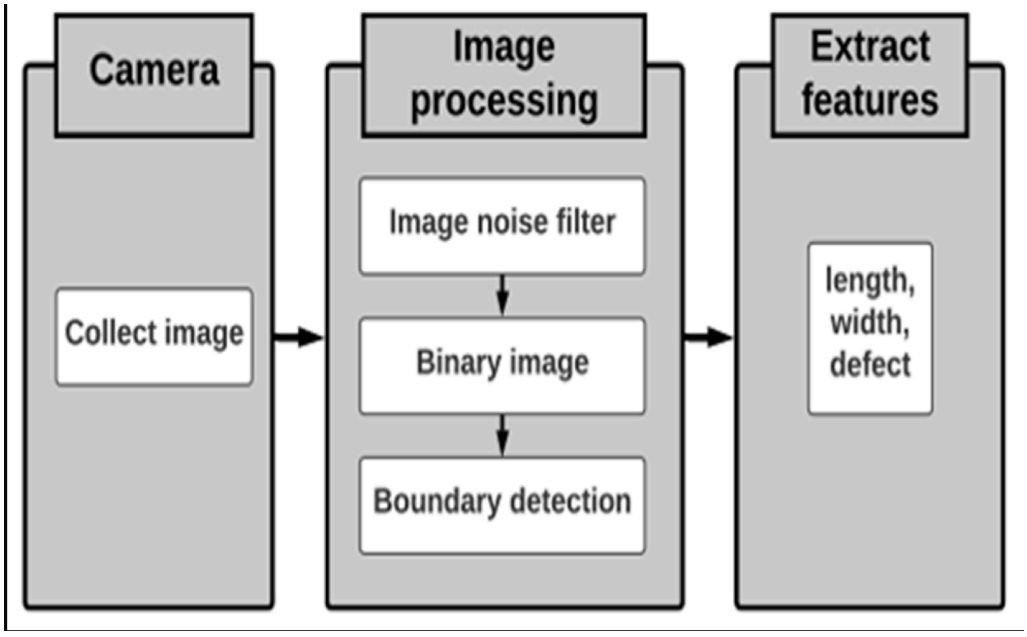
1. Enhanced Export Quality: By accurately predicting mango quality attributes such as ripeness, flavor, and shelf life, exporters can ensure that only high-quality fruits are selected for export. This leads to improved customer satisfaction and strengthens the reputation of the exporting company in international markets.
2. Market Differentiation: Implementing a machine learning-based quality prediction system sets exporters apart from competitors who rely on traditional methods. The ability to offer

precise quality assessments gives exporters a competitive edge and positions them as leaders in the industry.

Concept Generation:

The mango dataset, including parameters like size, color, weight, and internal qualities such as sugar content and acidity, is imported into the Jupyter environment. ML models are trained to predict mango quality based on these features.

- In the mango sorting system, mangoes are rotated on a roller conveyor and captured by cameras to assess external features. They then move to a tray conveyor for weighing and sorting based on a central processing unit's signals..
- In the first step, the images are acquired through the roller conveyor system inside the image processing chamber which is sealed and lighted. In the second step, captured images are processed by multiple algorithms such as increased fps (frames per second), image noise filter, edge detection, and boundary tracking. In the final step, length, width, and defect are extracted to generate a dataset



- ML models are trained to predict mango quality based on these features. High-density mangoes (density > 1.0) are sweeter and higher quality, while lower-density mangoes (density < 1.0) are often sour and lower quality.

Final product prototype:
The system includes data collection, model training, and implementation stages to ensure accurate and consistent mango grading.

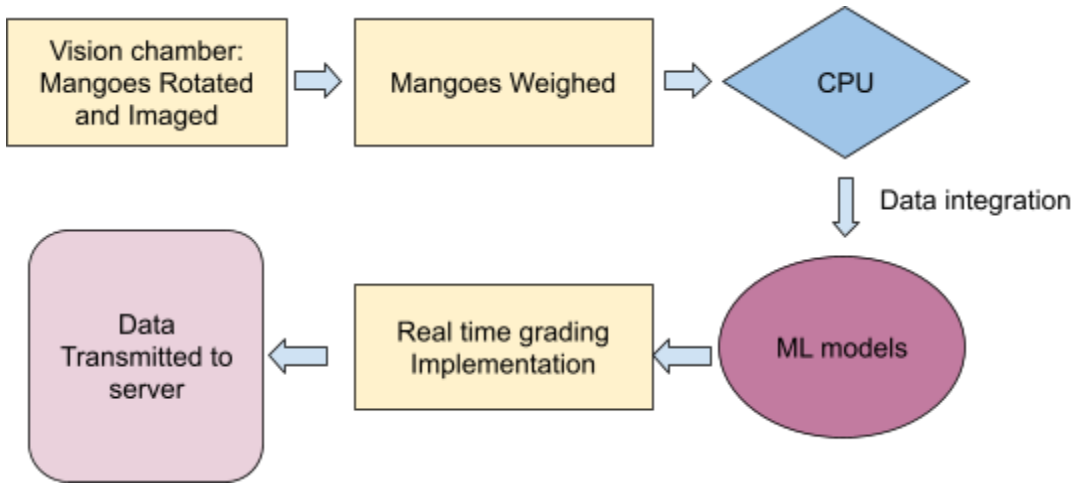
System Components

1. Conveyor and Vision System
    - Roller Conveyor: Transports mangoes into the vision chamber.
    - Cameras: Capture images from multiple angles as mangoes rotate on the rollers.
    - Tray Conveyor Weighs and sorts mangoes based on processing signals.

2. Central Processing Unit (CPU)
    - Data Integration Combines weight and external features data.
    - Data Structuring Converts unstructured data into structured format for ML models.

3. Machine Learning Model
    - Libraries: pandas, scikit-learn, Seaborn, Numpy for data handling and model training.
    - Training Data: Includes size, color, weight, sugar content, acidity, and density .
    - Models: LinearSVM, SVM, RandomForest, LogisticRegression

4. Data Transmission
    - Server: Ensures all processed data is securely transmitted and stored.
    - Deploy the ML model in the central processing system.
    - Real-time grading of mangoes as they pass through the conveyor system.
    - Continuous data transmission to the server for monitoring and analysis.



Product Details:
The product employs machine learning to assess mango quality for export by analyzing attributes like size, color, and aroma.

1. *Data Sources:* Comprehensive mango datasets and real-time sensor data from devices capture detailed fruit characteristics
2. *Algorithms, Frameworks, Software, etc. Needed:* -
    - ❖ Algorithm For each mango to be graded perform the following steps:

a. Step 1: Acquire a mango image using a proper image acquisition device.
b. Step 2: Perform median filtering and image resizing on the acquired image to obtain a preprocessed image.
c. Step 3: Convert the RGB image to grayscale and apply simple thresholding using a fixed threshold to obtain a binary image.
d. Step 4: Perform anding of binary image and preprocessed image to get the mango segmented image.
e. Step 5: Extract color, geometric, and shape features from mango segmented image.
f. Step 6: Determine ripeness, size, and shape using pre-trained classification models.
g. Step 7: Determine mango grade using simple decision-making or a pre-trained grade classifier.

❖ Intermediate Results and Discussion

Machine learning algorithms, data processing tools like TensorFlow, cloud computing services, and software development tools are essential.

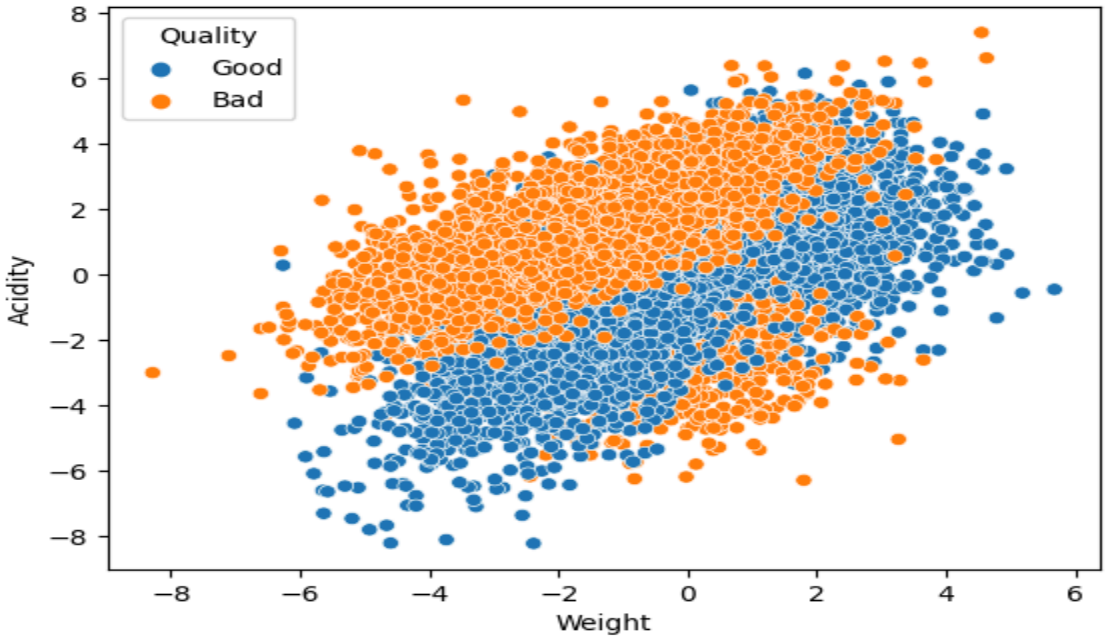
3. *Team Required to Develop*: Data scientists, software engineers, and agricultural experts collaborate to develop and deploy the predictive analytics platform.
4. *The cost* of developing such a system varies depending on factors like data acquisition methods, and infrastructure requirements. It includes expenses related to data collection, model development, software development, hardware procurement, and ongoing maintenance and support.

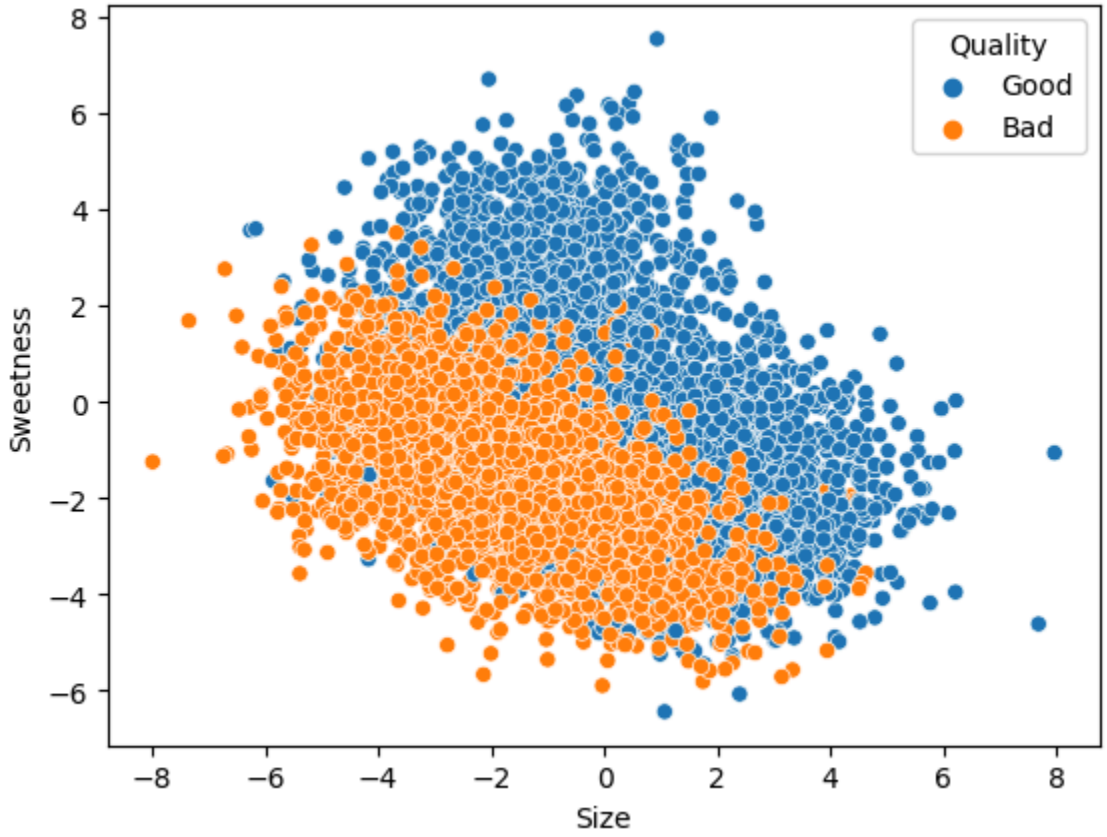
Code implementation on a similar dataset (available data is about banana fruit)**:**

conducted a code implementation on a similar dataset, focusing on banana fruit due to the unavailability of mango data. Utilizing logistic regression and SVM from the Scikit library, the SVM model yielded promising results, achieving an accuracy of 98.4%. You can access the code implementation on Kaggle via the following link: https://www.kaggle.com/code/dipalisabnis/banana-quality-prediction?scriptVersionId=177782714

Additionally, I generated visualizations from exploratory data analysis (EDA).

- From the first diagram, acidity increases with mango weight, impacting mango quality. Conversely, excessive acidity and weight lead to reduced quality, while moderate weight and acidity levels improve quality.
- In the second diagram, sweetness decreases as mango size increases. Balancing sweetness with size is crucial for determining mango quality."


Quality
Good
Bad
Acidity
Weight
8
6
4
2
0
−2
−4
−6
−8
−8
−6
−4
−2
0
2
4
6


Quality
Good
Bad
Sweetness
Size
8
6
4
2
0
−2
−4
−6
−8
−8
−6
−4
−2
0
2
4
6
8

The correlation between different features:

	Size	Weight	Sweetness	Softness	HarvestTime	Ripeness	Acidity
Size	1.000000	-0.181933	-0.258027	0.164526	0.582014	0.042587	-0.140673
Weight	-0.181933	1.000000	0.422683	-0.189908	-0.079465	-0.035276	0.443879
Sweetness	-0.258027	0.422683	1.000000	-0.095338	-0.201657	0.180095	0.187059
Softness	0.164526	-0.189908	-0.095338	1.000000	0.194734	-0.253804	-0.146014
HarvestTime	0.582014	-0.079465	-0.201657	0.194734	1.000000	0.107127	-0.090903
Ripeness	0.042587	-0.035276	0.180095	-0.253804	0.107127	1.000000	-0.351531
Acidity	-0.140673	0.443879	0.187059	-0.146014	-0.090903	-0.351531	1.000000

Conclusion:

Using machines to guess mango quality could help a lot. It might make things faster and better. But, we need everyone to work together. We also have to keep making the guessing better. If we keep improving, mango exports could get stronger and make more money.

In conclusion, the development of a machine learning-based system for predicting mango quality presents significant opportunities for improving efficiency and accuracy in the mango export industry. The data-driven insights and predictive modeling techniques, we can streamline the quality assessment process, enhance decision-making, and ultimately elevate the competitiveness of Indian mango exports on the global market. However, successful implementation will require collaboration across stakeholders, ongoing refinement of algorithms, and adherence to regulatory standards. With continued innovation and investment in technology, the future of mango quality prediction holds promise for driving sustainable growth and prosperity in the industry.

References:

1. ***C.S. Nandi, B. Tudu, and C. Koley,-Machine vision technique for grading of harvested mangoes based oN maturity and quality. IEEE Sensors Journal, 2016, vol.16(16), pp. 6387-6396. DOI10.1109/JSEN.2016.2580221***
2. ***www.google.com***
3. ***https://www.sciencedirect.com***
4. ***https://inpressco.com/wp-content/uploads/2021/02/Paper170808-812.pdf***
5. ***www.mdpi.com***
6. ***https://apeda.gov.in/apedawebsite/SubHead_Products/Mango.htm***