International Journal of Scientific Research and Engineering Development-– Volume 7 Issue 1, Jan-Feb 2024

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Machine Learning as a Strategic Tool: A Comprehensive Literature Review for Advancing Agricultural Analysis, with Emphasis on the Cocoa Bean Quality Assessment

Andrey M. Omas-as*, Jewelle Abbygale M. Daang**, Edwin R. Arboleda***

*(Department of Computer, Electronics, and Electrical Engineering, Cavite State University, Indang, Cavite, Philippines
Email: mrs15.omasasam@gmail.com)

** (Department of Computer, Electronics, and Electrical Engineering, Cavite State University, Indang, Cavite, Philippines
Email: mrs6.daangjam@gmail.com)

*** (Faculty of Department of Computer, Electronics, and Electrical Engineering, Cavite State University, Indang, Cavite, Philippines
Email: edwin.r.arboleda@cvsu.edu.ph)

--************************----------------------------------

Abstract:

As precision agriculture develops prominence for sustainable cacao production in the face of global challenges the role of machine learning (ML) as an essential tool in automating and optimizing cacao bean quality assessment becomes significant. The paper provides an in-depth review and comparative analysis of ML applications in cocoa bean quality assessment, incorporating insights from 20 relevant publications. Methodologies spanning from conventional algorithms to deep learning (DL) models and hybrid approaches are thoroughly studied. The discussion contextualizes these findings within the environment of cocoa bean quality assessment, comparing them to existing literature and addressing the implications and limitations of ML in this domain. Addressing unexpected or contradictory outcomes provides deep insights into the different aspects of ML for cocoa bean studies. This study contributes to a more detailed knowledge of the strengths and limitations of various ML techniques in cocoa bean quality assessment, which will guide future research for successful and sustainable cocoa production.

Keywords **—Precision Agriculture, Machine Learning (ML), Cocoa Bean Quality, Sustainable Production, Deep Learning (DL).**

--************************----------------------------------

I. INTRODUCTION

Cocoa is a crop that has been widely recognized as a plantation commodity [1], [2]. In recent decades, it has proven to be an effective source of income for farmers, leading to better living conditions [3], [4]. It has played a crucial role in promoting economic growth, particularly in rural production centers, and has contributed significantly to foreign exchange earnings [5], [6]. With over a million workers involved in its production, processing, and trade sectors, cocoa has become a vital part of the economy [5], [7]. The commitment to cocoa as a factor for development serves as a driving force behind the sustained progress in refining techniques for cocoa bean quality assessment, ensuring the ongoing improvement and sustainability of this crop. This sustained commitment reflects a holistic approach to bolstering not only economic aspects but also the overall resilience and viability of cocoa production systems [8], [9], [10], [11].

In line with this, the study investigates the use of machine learning (ML) in agricultural practices, specifically its potential for improving cocoa bean quality assessment. This paper examines prominent aspects of established ML models, explaining their contributions using a carefully selected collection of relevant literature. A list of abbreviations used is provided for better comprehension.

TABLE I
ABBREVIATIONS THAT APPEAR IN THESTUDY

Abbreviation	Expanded Form
ANFIS	Adaptive-Neuro Fuzzy Interference Systems
ANN	Artificial Neural Network
CART	Classification and Regression Detector
CNN	Convolutional Neural Network
CNN-SVM	Convolutional Neural Network with Support Vector Machine

DL	Deep Learning
DM	Dry Matter
DNN	Deep Neural Network
DT	Decision Tree
E-nose	Electronic Nose
FI	Fermentation Index
FmD	Fermentation Duration
GA	Genetic Algorithm
GDI	Gini Diversity Index
GLCM	Gray Level Co-occurrence Matrix
IoT	Internet of Things
KNN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
Mc	Moisture content
MaxEnt	Maximum Entropy
MELS-SVM	Multiclass Ensemble Least-Squares Support Vector Machine
ML	Machine Learning
MLP-ANN	Multi-Layer Perceptron with Artificial Neural Network
MOS	Metal Oxide Semiconductor
MVA	Multivariate Analysis
NB	Naïve Bayes
NIRS	Near-Infrared Spectroscopy
PCA	Principle Component Analysis
PCA-SVM	Principle Component Analysis with Support Vector Machine
PLSR	Partial Least Squares Regression
RF	Random Forest or Bootstrap Forest
RMSE	Root Mean Square Error
SVM	Support Vector Machine
TA	Titratable Acidity
VG-16	Visual Geometry Group Network

ML has provided new prospects for data-driven research and quick decision-making in agricultural analytics [12], [13], [14]. However, ML adoption in agriculture confronts hurdles such as data issues and potential misuse of evaluation methods [15], [16], [17], [18]. Despite these challenges, ML has proven to be effective in many areas of agriculture, including crop, livestock, water, and soil management, with a particular emphasis on crop management [19], [20]. ML algorithms provide timely analytical views to assist in proactive, data-driven decision-making in the agri-food supply chain [12], [20], [21]. ML is being used to model many aspects of agricultural production systems, providing significant information for farm-level management decisions [12], [19], [22]. Various ML approaches, including ANN, have been applied in these applications [14], [23], [24], [25], [26].

Numerous techniques have been put forth to evaluate the quality of cocoa beans. Some examples are categorizing cocoa beans according to their spectral fingerprints, especially during later stages of fermentation, Bayona (2018) proposes using closed range hyperspectral pictures [27]. Santika (2018) determines the export grade of cocoa beans using a hybrid ANFIS and GA [28]. In contrast to conventional classification techniques, Yro (2018) used a machine vision system combined with an SVM model to quickly, accurately, and consistently distinguish cocoa beans based on the degree of fermentation [29]. When Banboye (2020) compares several greenhouse technologies to analyze the drying behavior and quality of cocoa beans, she discovers that a customized greenhouse dryer produces the greatest results [30]. Ouokam (2019) emphasizes the financial and technical obstacles that farmers experience in enhancing the quality of their cocoa beans while highlighting the significance of agricultural techniques such as appropriate fermentation and drying [10]. Teye (2020) utilized NIRS to effectively prove the postharvest cocoa bean sector, particularly for qualitative and quantitative analysis of cocoa bean products [31].

This literature review aims to investigate the recent advancements of ML applications in the classification of cocoa bean quality. Specifically, it aims:

- To examine the state of ML in the cocoa bean industry today, highlighting its contribution to data-driven decision-making in a range of fields.
- To analyze current approaches and obstacles in utilizing ML for cocoa bean quality assessment, providing insights into its effectiveness and prospective improvements.
- To integrate research results to offer useful insights, tackle obstacles, and propose future paths for utilizing ML as a tactical instrument in agricultural evaluation, with an emphasis on evaluating cocoa bean quality.

The following parts could be fostered by integrating ML into the assessment of cocoa bean quality, which holds great potential to transform the cocoa industry:

- Intelligent Decision-Making. The literature study gives insights into the incorporation of ML in cocoa bean quality evaluation, arming stakeholders with the knowledge they need to make informed decisions utilizing advanced analytical approaches.
- Improved Agricultural Techniques. A better understanding of the role that ML plays in evaluating the quality of cocoa leads to improved agricultural practices, which in turn produces more effective and efficient quality control systems for the cocoa business.
- Quality Enhancement. The review demonstrates the potential for ML to improve cocoa bean quality evaluation by synthesizing pertinent literature, providing opportunities for improvement in manufacturing processes and end-product quality.
- Resource Optimization. By assuring optimal resource allocation, cutting waste, and optimizing quality evaluation procedures, ML integration can increase overall efficiency in the cocoa business.
- Future Innovation. The review lays the groundwork for future developments in cocoa evaluation, directing researchers, industry experts, and policymakers toward strategic ML adoption for long-term gains in cocoa production and quality management.

II. METHODS

This review was conducted in accordance with the guidelines of Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA)

A. *Search of Literature*

The literature search for this review adhered to a formal protocol outlined as follows.

B. *Criteria for Selecting Relevant Literature on ML in Agriculture*

To comprehensively identify literature related to ML applications in agriculture, particularly cocoa bean quality assessment, a systematic selection process was employed. Studies published between 2018 and 2024 were considered, focusing on recent developments in the field.

Explorations were performed on ResearchGate (https://www.researchgate.net/), IEEE Xplore, Scopus (www.scopus.com), Litmaps, and Google Scholar. The titles and abstracts of published studies were scrutinized using specific keywords, including "evaluation of cocoa bean quality," "utilization of ML," and "applications in agriculture." The Boolean operator ‘AND’ and appropriate wildcard symbols were employed for each database. Supplementary key terms like precision agriculture and DL were also included in the search strategy.

The inclusion criteria encompassed peer-reviewed articles, conference papers, and relevant reviews, with a specific emphasis on works demonstrating the integration of ML techniques for cocoa bean quality assessment. Additionally, the citation management tool Zotero was utilized to organize and streamline the selection process.
The researchers performed the initial review in the following steps:

1) *Identification:* The exploration involved the utilization of specific keywords such as "evaluation of cocoa bean quality," "utilization of ML," and "applications in agriculture." These databases were selected for their relevance to the fields of agriculture and technology.
2) *Screening:* Studies not aligned with cocoa bean quality assessment or those not published in English were excluded during this phase. The primary objective was to refine the initial set of records to focus on those most likely to meet the goals of the review.
3) *Eligibility:* Articles lacking clear methodologies or not demonstrating the integration of ML techniques for cocoa quality assessment were excluded. This stage ensured a detailed evaluation of potentially pertinent studies.
4) *Inclusion/Exclusion Criteria:* Inclusion criteria centered on peer-reviewed articles, conference papers, and reviews showcasing the integration of ML in cocoa quality assessment. Non-English publications and studies lacking clear methodologies were considered exclusion criteria. These criteria guided the selection of studies that align with the objectives of the review.

C. *Data Extraction*

The extraction of data from each related study within the study was conducted through a twofold process. Initially, the researchers analyzed two recent checklists to delineate the specific set of criteria relevant to the research. Following this, 20 related studies were examined to identify the most frequently reported data points across the included publications. This analysis aimed to formulate an extraction grid that would ensure a consistent, rigorous, and comprehensive approach to data extraction.

The actual data extraction process commenced subsequent to the initial phases. Each study underwent analysis for fundamental descriptive statistics, including quality assessment and reporting methods.

1) *Information Extraction:* Gather pertinent details from each chosen study, such as the title, authors,

publication year, ML techniques employed, key discoveries, and cocoa bean quality attributes evaluated.

2) *Accuracy Metrics:* Document reported accuracy rates and performance metrics from each study to compare the effectiveness of various ML methods.
3) *Study Constraints:* Identify and document any limitations mentioned in the selected studies, providing insights into potential challenges and areas for improvement in ML applications for cocoa bean quality assessment.

D. Methodological Approach for Chosen Studies

In examining studies related to the evaluation of cocoa bean quality through the application of ML, a thorough analysis of methodologies was systematically carried out. The selected studies underwent a detailed categorization process, considering various elements including:

1) *Study Selection Criteria:* Establish clear criteria for selecting studies that are consistent and relevant to the research objectives.
2) *Data Synthesis:* Summarize the main findings from each chosen study, evaluate them based on the ML techniques used and cocoa bean quality attributes evaluated.
3) *Comparative Analysis:* Conduct a comparative analysis of the reported accuracy rates, strengths, and limitations of different ML approaches to cocoa bean quality assessment.
4) *Implications and Recommendations:* Discuss the implications of the findings, identify common trends or challenges, and provide recommendations for future research in this domain.
5) *Detailed Methodological Analysis:* Provide a detailed examination of the ML algorithms utilized, methods for data collection and preprocessing, and techniques for feature selection and extraction across chosen studies.

This methodological approach facilitated a comprehensive exploration of the varied strategies utilized across the chosen studies, providing nuanced insights into the specific methodologies applied in the realm of cocoa bean quality assessment through ML.

III. RESULTS AND DISCUSSION

A. Presentation of Key Findings in MLApplications for Cocoa Bean Analysis

ML, a cutting-edge area of artificial intelligence, has completely changed the way computers analyze data, anticipate outcomes, and find patterns in enormous amounts of data. Systems can learn from experiences and gradually enhance performance thanks to their adaptive nature, which eliminates the need for explicit programming. ML has applications in many different disciplines, from predictive analytics to picture identification and natural language processing, from simple algorithms to intricate DL structures.

Many ML approaches have been effectively used to evaluate the quality of different agricultural goods, such as coffee beans, cocoa beans, and coconut sugar. Tan (2019) classified the degree of fermentation of cocoa beans using an E-nose system based on ML[32]. By presenting a hybrid CNN-SVM and PCA-SVM method for the classification of cocoa beans, Jean (2022) advanced the field and achieved great accuracy and robustness [33]. Together, these research demonstrate how ML can enhance the evaluation of agricultural products' quality.

This study collected and analyzed 20 accessible papers and compared each based on the ML techniques the proponents used, the summary of their paper and their accuracy rate.

B. Specific Outcomes Related to Cocoa Bean Quality Assessment

The study commenced by identifying 1,046 articles from diverse sources, such as IEEE, Google Scholar, and ResearchGate. Initial steps involved eliminating duplicates, resulting in a collection of 927 distinct studies. Subsequently, the titles were examined for relevance, leading to the exclusion of 378 articles. Difficulties in acquiring the complete texts for 98 studies left 271 for a detailed examination. Within this subset, 112 studies were discarded for not aligning with the designated topic, and 41 were excluded due to inadequate study design. In the end, 20 studies meeting all predefined criteria were incorporated into the conclusive analysis.

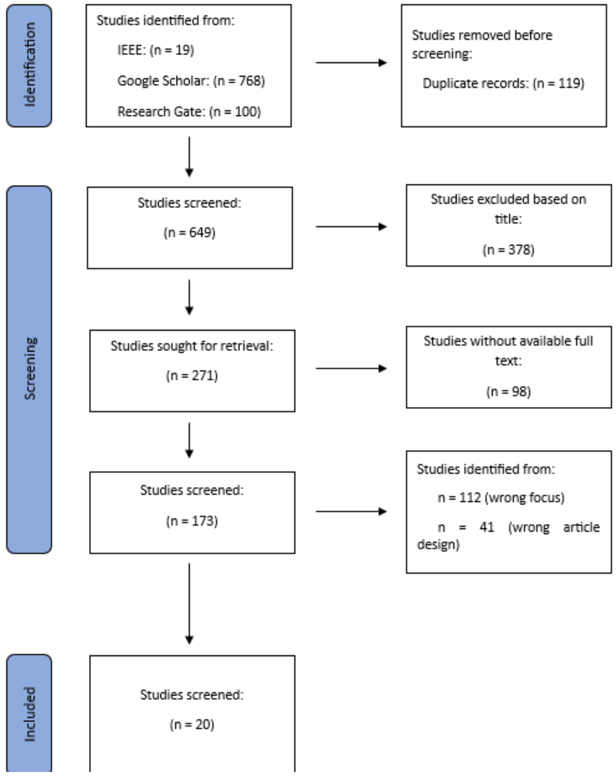

Fig 1. PRISMA flowchart for selection of publications

TABLE III
REVIEW OF INCLUDED STUDIES

Literature Title - Author (Year)	Models/ Algorithms	Summary	Result	Reference
Cacao Bean Quality Assessment Procedure: A Method for Classification Process - Brosas, D., Obediencia, D., & Villafuerte, R. (2018)	ANFIS	The research findings indicate that ANFIS proves to be an effective method for precisely classifying the quality levels of cacao beans.	Accuracy Rate: 99.715% Error Rate: 0.285%	[34]
Quality Evaluation of Fair-Trade Cocoa Beans from Different Origins Using Portable Near-Infrared Spectroscopy (NIRS) - Forte, M., Currò, S., Van de Walle, D., Dewettinck, K., Mirisola, M., Fasolato, L., & Carletti, P. (2022)	NIRS	The findings of this research suggest that the use of NIRS, especially with the improvements seen in portable NIR spectrometers, provides an advantageous method for quickly and non-invasively assessing the quality characteristics of cocoa beans.	DM: 94% accuracy Protein: 93% accuracy Ash: 90% accuracy Lipids: 83% accuracy TA: 86% accuracy pH: 88% accuracy	[35]
Analyzing the Accuracy of KNN-based Cacao Bean Grading System - Oraño, J., Padao, F., & Malangsa, R. (2019)	KNN	This implies that the KNN model proficiently automates the categorization of cacao beans, presenting a more effective and precise substitute for manual grading methods.	93.33% accuracy	[36]
Electronic Nose Coupled with Linear and Nonlinear Supervised Learning Methods for Rapid Discriminating Quality Grades of Superior Java Cacao Beans - Hidayat, S., Rusman, A., Julian, T., Triyana, K., Veloso, A., & Peres, A. (2019)	E-nose, LDA, SVM, and ANN.	The research showcases that utilizing an E-nose equipped with eight MOS gas sensors and a moisture-temperature sensor, especially through the E-nose-MLP-ANN procedure, can effectively and accurately classify three quality grades of superior Java cocoa beans, indicating its potential as an efficient quality control tool in the cocoa industry that requires minimal sample preparation.	Training Dataset: 99% accuracy External-Validation Dataset: 95% accuracy	[37]
Classifying Physical Morphology of Cocoa Beans Digital Images using Multiclass Ensemble Least-Squares Support Vector Machine - Lawi, A., & Adhitya, Y. (2018)	MELS-SVM	The main result of this study is that the MELS-SVM, using morphological features extracted from digital images, effectively categorizes cocoa beans into four groups (Normal, Broken, Fractured, and Skin Damaged) with high accuracy.	99.705% accuracy	[38]
Application of portable near infrared spectroscopy for classifying and quantifying cocoa bean quality parameters - Anyidoho, E.	Portable NIRS and MVA	The main results in this study highlight the importance of FmD, FI, pH, and Mc as crucial quality attributes for cocoa beans.	100% accuracy	[39]

K., Teye, E., Agbemafle, R., Amuah, C. L. Y., & Boadu, V. G. (2021)				
Application based on Hybrid CNN-SVM and PCA-SVM Approaches for Classification of Cocoa Beans - Ayikpa , K. J., Mamadou, D., Ballo, A. B., Gouton, P., & Adou, K. J. (2022)	CNN-SVM and PCA-SVM.	The study introduces a resilient hybrid CNN-SVM mode and PCA-SVM based model, showcasing enhanced efficacy in classifying cocoa beans post-harvest fermentation in comparison to conventional CNN and SVM classification approaches.	Hybrid CNN-SVM Model: 98.32% accuracy PCA-SVM Based Model: 97.65% accuracy	[33]
An intelligent cocoa quality testing framework based on deep learning techniques - Essah, R., Anand, D., & Singh, S. (2022)	CNN, DNN, and VG-16 models.	The results of this study contribute to the development and validation of a quality testing framework for cocoa beans, incorporating transfer learning with DL techniques like CNN and DNN, to offer a more objective and efficient assessment of post-harvest fermentation, implemented using Python in the Jupyter Notebook.	Transfer Learning Fine Tuning: 45.41% accuracy VG-16 (VGG-16): 43% accuracy	[40]
A Decision Tree Classification Model for Cocoa Beans Quality Evaluation based on Digital Imaging - Tanimowo, A. O., Eludiora, S., & Ajayi, A. (2020)	CART.	The study examines the use of a DT ML algorithm for quality classification of cocoa beans, the need of standardizing cocoa bean assessment, and the use of hyper-spectral imaging and an E-nose to monitor cocoa bean quality. It also explains how to capture cocoa bean photos, process them, extract features, and develop a prediction model for categorization. The article also compares classification results produced from various dividing criteria to those of government produce inspectors.	89.2% accuracy	[41]
Determination of Cocoa Bean Quality with Image Processing and Artificial Neural Network - Astika, I. W., Solahudin, M., Kurniawan, A., & Wulandari, Y. (2010)	ANN	The research demonstrates the effectiveness of a DT, a supervised ML approach, for classifying cocoa beans into categories such as good, slatty, mouldy, and weevilly based on digital images captured by a digital camera.	First ANN Structure: 84%, 52%, 20%, and 20% accuracy Second ANN Structure: 99%, 98%, 79% accuracy	[42]
Identifying the Quality System of Cocoa Beans to Increase Productivity Using Backpropagation Neural Network Algorithm: A Case Study at Sumber Rejeki Farmers Group, Patuk Gunung Kidul - Lestari, U., Kumalasanti, R. A., & Wulandari, E. (2019)	ANN Backpropagation	The study's main findings demonstrate the effectiveness of the Backpropagation method with specific parameters in ANN for accurately classifying cocoa beans into quality and non-quality categories based on image processing results.	Error Rate: 30% Accuracy Rate: 70%	[43]
Classification of Cacao Beans Based on their External Physical Features Using Convolutional Neural Network -	Pre-trained CNN	The key outcome of this study is the effective creation of a computer vision system that utilizes image processing and a pre-trained neural network	90.67% accuracy	[44]

Hortinela, C. C.; Tupas, K. J. R. (2022)		to precisely categorize cacao beans according to their external physical quality, addressing challenges associated with manual sorting.		
Cocoa beans classification using enhanced image feature extraction techniques and a regularized Artificial Neural Network model - Eric, O.; Mensah Gyening, R.-M. O.; Appiah, O.; Takyi, K.; Appiahene, P. (2023)	ANN	The primary conclusion drawn from this study is that the proposed color and texture extraction methodology, in conjunction with a generalized Neural Network model, markedly improves the machine classification of Cut-Test cocoa beans. This enhancement results in a more uniform classification rate across all grades, surpassing the performance of alternative ML models. Consequently, the developed approach stands as a dependable Computer-Aided Diagnostic tool for the precise classification of cocoa beans.	Accuracy: 85.36% Precision: 85% Recall: 83% F1 Measure: 83%	[45]
Application of hyperspectral imaging for cocoa bean grading with machine learning approaches -Liu, N.; Gonzalez, J. M.; Ottestad, S.; Hernandez, J. (2023)	MaxEnt. multiclass classification model	The study presents a non-invasive cocoa bean grading system using hyperspectral imaging and a maximum entropy multiclass classification model, demonstrating elevated accuracy in classifying crucial types of beans (good, under-fermented, and slaty) without the necessity of physically opening the beans.	Close to 80% accuracy	[46]
Quality Assessment Level of Quality of Cocoa Beans Export Quality using Hybrid Adaptive Neuro - Fuzzy Inference System (ANFIS) and Genetic Algorithm - Santika, G. D.; Wulandari, D. A. R.; Dewi, F. (2018)	ANFIS and GA	The research findings indicate that the application of the ANFIS method within an expert system, supported by GAs, results in an effective and accurate assessment of cocoa bean quality, as evidenced by a RMSE of 4.3. This underscores the feasibility of employing these algorithms as an expert system for the selection of cocoa bean quality.	RMSE is reported as 4.3	[28]
Feature Extraction for Cocoa Bean Digital Image Classification Prediction for Smart Farming Application -Adhitya, Y.; Prakosa, S. W.; Köppen, M.; Leu, J.-S. (2020)	GLCM and CNN	The research focuses on integrating Industry 4.0 principles into agriculture, specifically in evaluating and categorizing cocoa beans, utilizing artificial intelligence methods like GLCM and CNN for feature extraction. Findings reveal that GLCM texture extraction yields more dependable results than CNN. The approach, executed through on-site preprocessing on a low-performance computational device, aims to encourage the adoption of modern IoT technologies among farmers, ultimately enhancing the overall security of the food supply chain.	SVM Classifier: 59.14% accuracy XGBoost Classifier: 56.99% accuracy GLCM Textural Features with SVM Classifier: 61.04% accuracy GLCM Textural Features with XGBoost Classifier: 65.08% accuracy	[47]

Sensing fermentation degree of cocoa (Theobroma cacao L.) beans by machine learning classification models based electronic nose system - Tan, J.; Balasubramanian, B.; Sukha, D.; Ramkissoon, S.; Umaharan, P. (2019)	RF, Boosted tree, DT, ANN, NB, KNN.	The study introduces a ML-based E-nose system as a fast and affordable method to determine the fermentation degree of cocoa beans, comparing six ML methods. The RF algorithm achieved a low misclassification rate, while ANN and boosted tree demonstrated the effectiveness of the proposed method over other classification techniques for cocoa beans.	9.4% error rate	[32]
Cocoa Beans Digital Image Classification Based On Color Features using Multiclass Ensemble Least-Squares Support Vector Machine - Adhitya, Y., Lawi, A., Hartono, & Mursalin (2019)	MELS-SVM	The research involves utilizing image processing methods such as grayscale processing, thresholding, and edge detection, combined with color feature extraction, to categorize cocoa beans into three groups (Fermented, Un-Fermented, and Moldy)	Precision Level: 99.281%	[48]
Fermentation Level Classification of Cross Cut Cacao Beans Using k-NN Algorithm - Angelia, R. E.; Linsangan, N. B. (2018)	KNN	The research findings highlight the effectiveness of employing image processing and the KNNs Algorithm to classify cacao beans based on fermentation levels in chocolate production post-harvest procedures.	Well-fermented beans: 97.22% accuracy Under-fermented beans: 92.59% accuracy Over-fermented beans: 75% accuracy Analyzing unknowns: 80% accuracy	[49]
Quality Control of Commercial Cocoa Beans (Theobroma cacao L.) by Near-infrared Spectroscopy - Hashimoto, J. C., Lima, J. C., Celeghini, R. M. S., Nogueira, A. B., Efraim, P., Poppi, R. J., & Pallone, J. A. L. (2018)	PLSR	This study showcases the efficiency of NIRS as a swift and environmentally friendly technique for precise predictions of diverse chemical attributes, encompassing moisture, pH, acidity, fat, shell content, protein, total phenolic compounds, caffeine, and theobromine in commercial cocoa beans.	Coefficient of Determination (R2) Range: 0.67 to 0.89 Relative Errors: Below 10.2%	[50]

C. Comparative Analysis of Different MLApproaches in this Context

The accuracy rates demonstrate the effectiveness of several ML approaches in a variety of applications. Among the supervised learning methods, ANFIS achieves an outstanding 99.72% accuracy. The accuracy rates of KNNs and SVMs are 93.33% and 99%, respectively. ANNs in various topologies produce outcomes ranging from 20% to 99%, underlining neural networks' sensitivity to architecture and input.

CNNs demonstrate the efficacy of DL models with a high accuracy of 90.67%. Furthermore, when used in conjunction with PCA and SVMs, ANNs and CNNs exhibit competitive accuracy rates, underscoring the adaptability of these architectures in identifying intricate patterns.

However, the stated accuracy rates also highlight difficulties and discrepancies. Fine-tuning shows a reduced accuracy of 45.41%, indicating possible convergence issues, especially when combined with a learning rate. Comparably, using VG-16 results in a lower accuracy of 43.96%, highlighting the significance of selecting the right model and fine-tuning its parameters.

Ensemble approaches including RF, Boosted Tree, and DT, as well as standard ML algorithms like Nave Bayes and KNNs, routinely outperform, with a combined accuracy rate of 90.6%. These methods stress ensemble techniques' dependability and robustness in capturing complicated relationships in data.

The comparison also covers domain-specific techniques such as NIRS and GA. NIRS achieves accuracy rates ranging from 83% to a perfect 100% when used in a benchtop setup and as standalone technology, indicating its usefulness in various fields. However, when paired with ANFIS, the GA does not produce a reported accuracy, highlighting potential issues or limitations in the implementation of this specific technology.

To summarize, the varying accuracy rates observed in ML techniques highlight the significance of carefully considering the type of data, job complexity, and domain-specific needs when choosing the optimal solution. The published results contribute to the continuous investigation and improvement of ML approaches by providing insightful information about the advantages and drawbacks of each strategy.

D. Analysis of Findings within the Context of Cocoa Bean Quality Assessment

The findings presented in the study demonstrate the effectiveness of ML techniques in evaluating cocoa bean quality. The studies encompass a range of approaches, from traditional algorithms like KNNs and SVMs to advanced DL models such as CNNs and hybrid strategies like CNN-SVM and PCA-SVM.

The ANFIS achieves an impressive accuracy rate of 99.72%, highlighting the robustness of fuzzy logic-based models in accurately categorizing cocoa bean quality levels. On the other hand, the DT Classification Model achieves an accuracy of 89.2%, showcasing the effectiveness of simpler algorithms in quality assessment.

The studies also demonstrate the diversity of data sources employed for cocoa bean assessment, including E-nose systems, NIRS, and hyperspectral imaging. These ML approaches can evaluate various quality attributes, such as fermentation degree, Mc, and chemical composition, showcasing their holistic nature in capturing multiple aspects of cocoa bean quality.

E. Comparisons with Existing Literature on Cocoa Quality Analysis and ML

The results of the studies demonstrate that ML has become a versatile tool for cocoa quality analysis. The studies collectively show that different ML techniques can adapt to specific needs within the cocoa industry. For instance, Tan (2019) introduced an E-nose system, demonstrating its efficacy in determining cocoa bean fermentation degree [32].

The consistent effectiveness of NIRS across studies, as demonstrated by Forte et al. (2022) and Anyidoho et al. (2021), indicates the reliability and widespread applicability of this technique in assessing various quality parameters [35], [39].

The emergence of hybrid models, such as CNN-SVM and PCA-SVM, reflects the trend toward amalgamating the strengths of different algorithms to achieve heightened accuracy and robustness. Ayikpa et al.'s (2022) demonstration of the effectiveness of hybrid models in classifying cocoa beans post-harvest fermentation exemplifies this trend [33].

F. Exploration of the Implications and Limitations of ML in this Specific Application

While ML has shown promising results in evaluating cocoa bean quality, it is important to recognize its limitations and potential drawbacks. The fluctuation in accuracy rates across diverse ML models underscores the sensitivity of these methods to various factors, including data quality, preprocessing techniques, and model architecture.

The study by Essah et al. (2022), utilizing DL techniques, reports lower accuracy rates (45.41% and 43.96%) for Transfer Learning Fine Tuning and VG-16 models. This signals challenges in achieving optimal performance, emphasizing the necessity for meticulous model selection and parameter tuning [40].

Moreover, the diversity in methods and data sources poses challenges for standardization. Different studies focus on various quality attributes, making it arduous to establish a universal framework for cocoa bean quality assessment. Additionally, the requirement for specialized equipment, such as E-nose systems and hyperspectral imaging devices, may constrain the accessibility of these methods in certain regions or for smaller-scale producers.

G. Addressing any Unexpected or Divergent Findings in Cocoa Bean Quality Assessment

The variable accuracy rates across different models are the unexpected findings in ML for cocoa bean quality assessment. For instance, the study by Adhitya et al. (2020) suggests that GLCM texture extraction produces more reliable results than CNN, challenging the prevailing belief that DL consistently outperforms traditional feature extraction methods [48].

Furthermore, the lower accuracy rates observed in certain DL models, as seen in Essah et al. (2022), indicate that more intricate architectures do not always guarantee superior performance [40]. This calls for a nuanced approach in selecting and fine-tuning models based on the specific characteristics of cocoa bean data.

The exploration of ML applications in cocoa bean quality assessment provides valuable insights into the strengths and limitations of various techniques. As the field continues to evolve, addressing these limitations and refining methodologies will contribute to the development of more robust and universally applicable models for cocoa quality analysis.

IV. CONCLUSIONS

A. Summarization of Key Findings in ML for Cocoa Bean Quality Assessment

In conclusion, the broad variety of ML approaches analyzed demonstrates differing accuracy rates across various applications. The ANFIS has an amazing accuracy of 99.72%, although standard methods such as KNNs and SVMs perform well. DL models, notably CNNs, achieve 90.67% accuracy. However, obstacles remain, as seen by poorer fine-tuning accuracy rates and the VG-16. Ensemble approaches and classical algorithms operate well together, with an accuracy rate of 90.6%. The comparison highlights the necessity of picking the best suited strategy based on individual data characteristics and application needs, providing significant insights for ML methodology optimization.

B. Overall Interpretation and Significance of the Study

A wide range of performance across multiple applications is shown by a thorough evaluation of ML approaches. With an impressive accuracy of 99.72%, ANFIS stands out as a top performer, demonstrating its effectiveness in particular domains. Conventional algorithms, such SVM and KNNs, perform well and consistently show their dependability on a variety of applications. CNNs, one of the most popular DL models, show good accuracy; nonetheless, there are noticeable issues with fine-tuning and specific network topologies. Both ensemble techniques and classical algorithms provide consistent performance overall, offering dependable substitutes. The importance is in the nuanced comprehension of the advantages and disadvantages of each technique, which helps practitioners select the best strategy based on the particulars of the task at hand and the complexities of the dataset. This comparison provides useful insights into the continuing refining and refinement of ML approaches for a variety of applications.

C. Implications for Future Research in MLApplications for Cocoa Quality

The comparison highlights the necessity for continuing and nuanced research of these strategies in the context of cocoa quality assessment due to the different performance of ML algorithms in various applications. ANFIS's outstanding accuracy shows that investigating hybrid models, such as merging fuzzy logic with neural networks, may offer potential for improving cocoa quality evaluation. Traditional algorithms, such as KNNs and SVMs, are also robust, indicating that simpler approaches should not be neglected. In some circumstances, the success of CNNs demonstrates the possibility of employing DL for image-based cocoa quality assessments. Future study could focus on enhancing DL architectures, taking into account elements such as fine-tuning and network topology to improve their performance. Reliability is demonstrated by ensemble approaches like Boosted Tree and RF, which point to the need for more research on ensemble techniques to increase accuracy. Furthermore, given the domain-specific character of cocoa quality evaluation, further research into the combination of GAs with NIRS may be necessary. To sum up, future studies should concentrate on a comprehensive comprehension of the advantages and disadvantages of various ML approaches with the goal of creating hybrid models that are suited to the complexities of cocoa quality evaluation for effective and sustainable cocoa production.

D. Closing Remarks on the Study's Contribution to Advancing Cocoa Bean Quality Assessment Using ML.

Finally, the comprehensive analysis of various ML techniques in this study not only sheds light on the diverse landscape of cocoa bean quality assessment, but it also provides useful information that could substantially improve the topic by guiding the development of tailored and effective models for sustainable cocoa production.

ACKNOWLEDGMENT

The authors would like to express their gratitude to their respected academic advisors and mentors from the Department of Computer and Electronics Engineering at Cavite State University. Their guidance and support played a crucial role in shaping the literature review titled "Machine Learning as a Strategic Tool: A Comprehensive Literature Review for Advancing Agricultural Analysis, with Emphasis on the Cocoa Bean Quality Assessment." The authors are grateful to the pioneering researchers and scholars in the field, whose significant contributions laid the groundwork for their study. They extend special recognition to Cavite State University for supplying the necessary resources, and to their peers for participating in insightful discussions that enhanced the study. Finally, the authors convey their sincere gratitude to their families and friends for their constant support throughout this study.

REFERENCES

[1] K. Ruslan and O. Prasetyo, "Plantation Crop Productivity: Coffee, Sugarcane and Cocoa," Center for Indonesian Policy Studies, Jakarta, Indonesia, 2021. doi: 10.35497/349327.

[2] A. Venturieri et al., "The Sustainable Expansion of the Cocoa Crop in the State of Pará and Its Contribution to Altered Areas Recovery and Fire Reduction," JGIS, vol. 14, no. 03, pp. 294–313, 2022, doi: 10.4236/jgis.2022.143016.

[3] S. Jumiyati, "INCREASING INCOME OF COCOA FARMING THROUGH THE ROLE OF AGRICULTURAL EXTENSION AND STRENGTHENING INSTITUTIONAL CAPACITY OF FARMERS," IJAEB, vol. 04, no. 06, pp. 110–121, 2019, doi: 10.35410/IJAEB.2019.4465.

[4] H. E. Hernández-Núñez et al., "Cacao cultivation as a livelihood strategy: contributions to the well-being of Colombian rural households," Agric Hum Values, vol. 39, no. 1, pp. 201–216, Mar. 2022, doi: 10.1007/s10460-021-10240-y.

[5] K. Tsowou and S. K. Gayi, "Trade Reforms and Integration of Cocoa Farmers into World Markets: Evidence from African and non-African Countries:," JAT, vol. 6, no. 1–2, p. 16, 2019, doi: 10.2991/jat.k.190916.001.

[6] I. T. A. R. Indah Sari Pananrangi, "Cocoa Export Performance to Economic Growth in South Sulawesi," pae, vol. 58, no. 1, pp. 440–448, Jan. 2021, doi: 10.17762/pae.v58i1.793.

[7] S. Hadinata and M. M. Marianti, "Analisis Dampak Hilirisasi Industri Kakao di Indonesia," JAM, vol. 12, no. 1, pp. 99–108, May 2020, doi: 10.28932/jam.v12i1.2287.

[8] P. Boeckx, M. Bauters, and K. Dewettinck, "Poverty and climate change challenges for sustainable intensification of cocoa systems," Current Opinion in Environmental Sustainability, vol. 47, pp. 106–111, Dec. 2020, doi: 10.1016/j.cosust.2020.10.012.

[9] C. M. Miranda, R. Moreno, and P. Moreno, "Protected-Denomination-of-Origin Cocoa Bean: Chain governance and Sustainability Performance," Economia agro-alimentare, no. 1, pp. 1–24, Jul. 2020, doi: 10.3280/ecag1-2020oa10066.

[10] E. W. Ouokam, Y. Yunxian, M. Y. James, M. O. Appiah, and G. Motlehewa, "Assessment of agricultural practices for improving quality of cocoa beans: South-West Cameroon," IJEAB, vol. 4, no. 3, pp. 767–777, 2019, doi: 10.22161/ijeab/4.3.25.

[11] Muhardi et al., "Sustainability of cocoa production in Indonesia," Aust J Crop Sci, no. 14(6):2020, pp. 997–1003, Jun. 2020, doi: 10.21475/ajcs.20.14.06.p2510.

[12] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine Learning in Agriculture: A Review," Sensors, vol. 18, no. 8, p. 2674, Aug. 2018, doi: 10.3390/s18082674.

[13] N. Tantalaki, S. Souravlas, and M. Roumeliotis, "Data-Driven Decision Making in Precision Agriculture: The Rise of Big Data in Agricultural Systems," Journal of Agricultural & Food Information, vol. 20, no. 4, pp. 344–380, Oct. 2019, doi: 10.1080/10496505.2019.1638264.

[14] N. Yadav, S. M. Alfayeed, and A. Wadhawan, "MACHINE LEARNING IN AGRICULTURE: TECHNIQUES AND APPLICATIONS," IJEAST, vol. 5, no. 7, pp. 118–122, Nov. 2020, doi: 10.33564/IJEAST.2020.v05i07.018.

[15] S. Condran, M. Bewong, M. Z. Islam, L. Maphosa, and L. Zheng, "Machine Learning in Precision Agriculture: A Survey on Trends, Applications and Evaluations Over Two Decades," IEEE Access, vol. 10, pp. 73786–73803, 2022, doi: 10.1109/ACCESS.2022.3188649.

[16] A. Cravero and S. Sepúlveda, "Use and Adaptations of Machine Learning in Big Data—Applications in Real Cases in Agriculture," Electronics, vol. 10, no. 5, p. 552, Feb. 2021, doi: 10.3390/electronics10050552.

[17] M. Cockburn, "Review: Application and Prospective Discussion of Machine Learning for the Management of Dairy Farms," Animals, vol. 10, no. 9, p. 1690, Sep. 2020, doi: 10.3390/ani10091690.

[18] H. F. Williamson et al., "Data management challenges for artificial intelligence in plant and agricultural research," F1000Res, vol. 10, p. 324, Apr. 2021, doi: 10.12688/f1000research.52204.1.

[19] L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris, and D. Bochtis, "Machine Learning in Agriculture: A Comprehensive Updated Review," Sensors, vol. 21, no. 11, p. 3758, May 2021, doi: 10.3390/s21113758.

[20] A. Sharma, A. Jain, P. Gupta, and V. Chowdary, "Machine Learning Applications for Precision Agriculture: A Comprehensive Review," IEEE Access, vol. 9, pp. 4843–4873, 2021, doi: 10.1109/ACCESS.2020.3048415.

[21] I. Santoso, M. Purnomo, A. A. Sulianto, and A. Choirun, "Machine learning application for sustainable agri-food supply chain performance: a review," IOP Conf. Ser.: Earth Environ. Sci., vol. 924, no. 1, p. 012059, Nov. 2021, doi: 10.1088/1755-1315/924/1/012059.

[22] J. Zhang, J. Liu, Y. Chen, X. Feng, and Z. Sun, "Knowledge Mapping of Machine Learning Approaches Applied in Agricultural Management—A Scientometric Review with CiteSpace," Sustainability, vol. 13, no. 14, p. 7662, Jul. 2021, doi: 10.3390/su13147662.

[23] P. Saikia, R. D. Baruah, S. K. Singh, and P. K. Chaudhuri, "Artificial Neural Networks in the domain of reservoir characterization: A review from shallow to deep models," Computers & Geosciences, vol. 135, p. 104357, Feb. 2020, doi: 10.1016/j.cageo.2019.104357.

[24] U. M. R. Paturi and S. Cheruku, "Application and performance of machine learning techniques in manufacturing sector from the past two decades: A review," Materials Today: Proceedings, vol. 38, pp. 2392–2401, 2021, doi: 10.1016/j.matpr.2020.07.209.

[25] S. Damilola, "A Review of Unsupervised Artificial Neural Networks with Applications," IJCA, vol. 181, no. 40, pp. 22–26, Feb. 2019, doi: 10.5120/ijca2019918425.

[26] T. Kamatchi and B. V. Kumar, "Application of Artificial Neural Network with Various Algorithms Tools in Numerous Sectors – A Review," pd, vol. 91, no. 4, Apr. 2022, doi: 10.37896/pd91.4/91411.

[27] O. Bayona, D. Ochoa, R. Criollo, J. Cevallos-Cevallos, and W. Liao, "Cocoa bean quality assessment using closed range hyperspectral images," in 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Honolulu, HI, USA: IEEE, Nov. 2018, pp. 622–626. doi: 10.23919/APSIPA.2018.8659490.

[28] G. D. Santika, D. A. R. Wulandari, and F. Dewi, "Quality Assessment Level of Quality of Cocoa Beans Export Quality using Hybrid Adaptive Neuro - Fuzzy Inference System (ANFIS) and Genetic Algorithm," in 2018 International Conference on Electrical

Engineering and Computer Science (ICECOS), Pangkal Pinang: IEEE, Oct. 2018, pp. 195–200. doi: 10.1109/ICECOS.2018.8605189.

[29] A. Yro, C. N'Zi, and K. Kpalma, "Cocoa Beans Fermentation Degree Assessment For Quality Control Using Machine Vision and Multiclass SVM Classifier," International Journal of Innovation and Applied Studies, Nov. 2018, Accessed: Jan. 12, 2024. [Online]. Available: https://www.semanticscholar.org/paper/Cocoa-Beans-Fermentation-Degree-Assessment-For-and-Yro-N'Zi/a4b9c70ffbb7bc1da3a9007a2a6e3acfa263af78

[30] F. D. Banboye, M. N. Ngwabie, S. A. Eneighe, and D. B. Nde, "Assessment of greenhouse technologies on the drying behavior of cocoa beans," Food Science & Nutrition, vol. 8, no. 6, pp. 2748–2757, Jun. 2020, doi: 10.1002/fsn3.1565.

[31] E. Teye, E. Anyidoho, R. Agbemafle, L. K. Sam-Amoah, and C. Elliott, "Cocoa bean and cocoa bean products quality evaluation by NIR spectroscopy and chemometrics: A review," Infrared Physics & Technology, vol. 104, p. 103127, Jan. 2020, doi: 10.1016/j.infrared.2019.103127.

[32] J. Tan, B. Balasubramanian, D. Sukha, S. Ramkissoon, and P. Umaharan, "Sensing fermentation degree of cocoa (Theobroma cacao L.) beans by machine learning classification models based electronic nose system," J Food Process Engineering, vol. 42, no. 6, p. e13175, Oct. 2019, doi: 10.1111/jfpe.13175.

[33] A. K. Jean, M. Diarra, B. A. Bakary, G. Pierre, and A. K. Jérôme, "Application based on Hybrid CNN-SVM and PCA-SVM Approaches for Classification of Cocoa Beans," IJACSA, vol. 13, no. 9, 2022, doi: 10.14569/IJACSA.2022.0130927.

[34] D. Obediencia, D. Brosas, and R. Villafuerte, "(PDF) Cacao Bean Quality Assessment Procedure: A Method for Classification Process." Accessed: Jan. 10, 2024. [Online]. Available: https://www.researchgate.net/publication/330194037_Cacao_Bean_Quality_Assessment_Procedure_A_Method_for_Classification_Process

[35] M. Forte et al., "Quality Evaluation of Fair-Trade Cocoa Beans from Different Origins Using Portable Near-Infrared Spectroscopy (NIRS)," Foods, vol. 12, no. 1, p. 4, Dec. 2022, doi: 10.3390/foods12010004.

[36] J. F. V. Orano, F. R. F. Padao, and R. D. Malangsa, "Analyzing the Accuracy of KNN-based Cacao Bean Grading System," in 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Laoag, Philippines: IEEE, Nov. 2019, pp. 1–5. doi: 10.1109/HNICEM48295.2019.9072790.

[37] Universitas Gadjah Mada et al., "Electronic Nose Coupled with Linear and Nonlinear Supervised Learning Methods for Rapid Discriminating Quality Grades of Superior Java Cocoa Beans," IJIES, vol. 12, no. 6, pp. 167–176, Dec. 2019, doi: 10.22266/ijies2019.1231.16.

[38] A. Lawi and Y. Adhitya, "Classifying Physical Morphology of Cocoa Beans Digital Images using Multiclass Ensemble Least-Squares Support Vector Machine," J. Phys.: Conf. Ser., vol. 979, p. 012029, Mar. 2018, doi: 10.1088/1742-6596/979/1/012029.

[39] E. K. Anyidoho, E. Teye, R. Agbemafle, C. L. Y. Amuah, and V. G. Boadu, "Application of portable near infrared spectroscopy for classifying and quantifying cocoa bean quality parameters," J Food Process Preserv, vol. 45, no. 5, May 2021, doi: 10.1111/jfpp.15445.

[40] R. Essah, D. Anand, and S. Singh, "An intelligent cocoa quality testing framework based on deep learning techniques," Measurement: Sensors, vol. 24, p. 100466, Dec. 2022, doi: 10.1016/j.measen.2022.100466.

[41] A. O. Tanimowo, S. Eludiora, and A. Ajayi, "A Decision Tree Classification Model for Cocoa Beans Quality Evaluation based on Digital Imaging," Ife Journal of Technology, vol. 27, no. 1, Art. no. 1, Dec. 2020.

[42] I. W. Astika, M. Solahudin, A. Kurniawan, and Y. Wulandari, "Determination of Cocoa Bean Quality with Image Processing and Artificial Neural Network," Computer Based Data Acquisition and Control in Agriculture, 2010, Accessed: Jan. 10, 2024. [Online]. Available: http://repository.ipb.ac.id/handle/123456789/41666

[43] U. Lestari, R. A. Kumalasanti, and E. S. Wulandari, "Identifying the Quality System of Cocoa Beans to Increase Productivity Using Backpropagation Neural Network Algorithm: A Case Study at Sumber Rejeki Farmers Group, Patuk Gunung Kidul," J. Phys.: Conf. Ser., vol. 1413, no. 1, p. 012033, Nov. 2019, doi: 10.1088/1742-6596/1413/1/012033.

[44] C. C. Hortinela and K. J. R. Tupas, "Classification of Cacao Beans Based on their External Physical Features Using Convolutional Neural Network," in 2022 IEEE Region 10 Symposium (TENSYMP), Mumbai, India: IEEE, Jul. 2022, pp. 1–5. doi: 10.1109/TENSYMP54529.2022.9864337.

[45] O. Eric, R.-M. O. M. Gyening, O. Appiah, K. Takyi, and P. Appiahene, "Cocoa beans classification using enhanced image feature extraction techniques and a regularized Artificial Neural Network model," Engineering Applications of Artificial Intelligence, vol. 125, p. 106736, Oct. 2023, doi: 10.1016/j.engappai.2023.106736.

[46] N. Liu, J. M. Gonzalez, S. Ottestad, and J. E. Hernández, "Application of hyperspectral imaging for cocoa bean grading with machine learning approaches," in Hyperspectral Imaging and Applications II, N. J. Barnett, A. A. Gowen, and H. Liang, Eds., Birmingham, United Kingdom: SPIE, Jan. 2023, p. 6. doi: 10.1117/12.2645888.

[47] Y. Adhitya, S. W. Prakosa, M. Köppen, and J.-S. Leu, "Feature Extraction for Cocoa Bean Digital Image Classification Prediction for Smart Farming Application," Agronomy, vol. 10, no. 11, p. 1642, Oct. 2020, doi: 10.3390/agronomy10111642.

[48] Y. Adhitya, A. Lawi, H. Hartono, and M. Mursalin, "Cocoa Beans Digital Image Classification Based On Color Features using Multiclass Ensemble Least-Squares Support Vector Machine," in Proceedings of the Proceedings of the 1st Workshop on Multidisciplinary and Its Applications Part 1, WMA-01 2018, 19-20 January 2018, Aceh, Indonesia, Aceh, Padang, Jakarta, Indonesia: EAI, 2019. doi: 10.4108/eai.20-1-2018.2281929.

[49] R. E. Angelia and N. B. Linsangan, "Fermentation Level Classification of Cross Cut Cacao Beans Using k-NN Algorithm," in Proceedings of the 2018 5th International Conference on Bioinformatics Research and Applications, Hong Kong Hong Kong: ACM, Dec. 2018, pp. 64–68. doi: 10.1145/3309129.3309142.

[50] J. C. Hashimoto et al., "Quality Control of Commercial Cocoa Beans (Theobroma cacao L.) by Near-infrared Spectroscopy," Food Anal. Methods, vol. 11, no. 5, pp. 1510–1517, May 2018, doi: 10.1007/s12161-017-1137-2.

