JUST AGRICULTURE
multidisciplinary e-Newsletter

Seed Quality Determination through Machine Learning

Amrutha Raghavan[1], Gurubasavaswamy B M[2], D Vijay Kumar Reddy[3], Gokulnath G[4*], K J Sivasangari[5]

[1] Indira Gandhi Krishi Vigyan Kendra, Raipur, Chattisgarh.
[2&3] University of Agricultural Sciences, Raichur, Karnataka.
[4*] AC & RI, Tamil Nadu Agricultural University, Madurai, Tamil Nadu.
[5] Kumaraguru Institute of Agriculture, Erode, Tamil Nadu.

ARTICLE ID: 52

Abstract

Seed quality plays a pivotal role in agricultural productivity, influencing germination rates, crop health, and yield. Traditional seed quality assessment methods often rely on manual inspection and laboratory-based analysis which can be time-consuming and prone to human error. Machine learning is subset of artificial intelligence (AI) offers a novel approach to seed quality determination by leveraging advanced algorithms to analyze large volumes of data. Through image processing, spectral analysis and evaluation of phenotypic traits, machine learning models can accurately assess parameters such as seed viability, size, shape and potential for germination. This technology enhances the speed, precision and efficiency of seed quality evaluation, enabling improved decision-making in agriculture. This explores the application of machine learning in seed quality determination, highlighting its benefits and potential to revolutionize modern agricultural practices.

Keywords: Machine Learning, Seed Quality, Artificial Intelligence, and Seed Viability.

Introduction

Seed quality is a critical determinant of crop yield, directly impacting agricultural productivity and food security. Traditionally, the evaluation of seed quality has relied on manual inspection and biochemical tests, which are labor-intensive, time-consuming and sometimes inconsistent due to human error. With the advent of digital agriculture, machine learning (ML) has emerged as a transformative technology for improving the efficiency and accuracy of seed quality determination.

To validate seed quality qualities including germination and vigor testing, quality assurance programs in the seed industry use a variety of techniques. These methods have limitations in terms of subjectivity, time consumption, and the destructive nature of seed quality

JUST AGRICULTURE
multidisciplinary e-Newsletter

evaluation. In fact, there is an increasing demand for effective techniques that can provide a fast, accurate, non-destructive, and objective way to determine the quality of seeds.

Loss of viability and vigor is typically associated with changes in the chemical composition and internal anatomical features of seeds, although these changes are difficult to identify visually. In the meanwhile, data on detailed characteristics linked to seed quality has been effectively gathered using approaches based on spectrometry and X-ray imaging techniques. For example, by obtaining a lot of spectral information, Fourier transform near-infrared (FT-NIR) spectroscopy has shown a lot of promise in the detection of seed chemicals.

Importance of Seed Quality in Agriculture

Seed quality encompasses various attributes such as germination rate, purity, vigor and health status, which collectively determine a seed's ability to produce a healthy and productive plant. High-quality seeds ensure better crop establishment, uniformity and resilience to diseases and pests. Conversely, poor-quality seeds can result in low germination rates, weak seedlings, and ultimately lower yields. The need for accurate and efficient seed quality assessment methods is critical.

Traditional methods for seed quality evaluation involve physical and physiological tests such as visual inspection, seed germination tests, tetrazolium testing and moisture content determination. While these methods are effective to some extent, they are subjective, labor-intensive and may not provide real-time results. Furthermore, variability in seed quality across different batches and environmental conditions necessitates a more robust and consistent approach, which machine learning can provide.

Machine Learning in Seed Quality Determination

Machine learning, a crucial subset of artificial intelligence (AI), focuses on creating algorithms that can learn from data, recognize patterns and make decisions with little to no human intervention. This technology is increasingly applied across various industries, including agriculture. In the context of seed quality assessment, machine learning offers a powerful tool for analyzing large dataset such as high-resolution images, spectral data and other phenotypic traits.

Machine learning models can process these complex data types to identify and evaluate key seed quality parameters such as size, color, shape, moisture content and even genetic traits. For example, models can be trained to detect damaged or diseased seeds by analyzing image

JUST AGRICULTURE
multidisciplinary e-Newsletter

patterns, while spectral analysis can be used to assess seed viability or germination potential. These AI-driven assessments provide a faster, more accurate and efficient way to ensure seed quality compared to traditional methods, enhancing the precision of modern agricultural practices.

Machine Learning Techniques for Seed Quality Analysis

Several machine learning techniques have been employed to enhance seed quality determination:

- ✓ **Image-based Analysis:** High-resolution images of seeds can be analyzed using ML algorithms to detect defects such as cracks, discoloration or deformities. Techniques like convolutional neural networks (CNNs) are particularly effective in image recognition tasks, enabling the identification of seed quality characteristics that are not visible to the naked eye.
- ✓ **Spectral Analysis:** Seeds emit unique spectral signatures when exposed to different wavelengths of light. Machine learning models can analyze these spectral signatures to determine seed composition, moisture content and the presence of contaminants or diseases. Methods such as support vector machines (SVMs) and random forests are often used in conjunction with spectral data to classify seed quality.
- ✓ **Predictive Modeling:** ML algorithms can predict seed germination rates, vigor and other quality parameters based on historical data. Regression models, decision trees and neural networks are some of the common techniques used to develop these predictive models.
- ✓ **Natural Language Processing (NLP):** NLP techniques can process textual data from seed labels, certificates, or quality reports to extract relevant information about seed quality. This can be particularly useful in automated seed testing laboratories where large volumes of text data need to be analyzed rapidly.

Applications of Machine Learning in Seed Quality Determination

Machine learning applications in seed quality assessment can be broadly categorized into the following:

- ✓ **Seed Germination Prediction:** Germination rate is a critical parameter for determining seed quality. Machine learning models, particularly deep learning approaches can predict germination rates by analyzing seed images, moisture content and other

JUST AGRICULTURE
multidisciplinary e-Newsletter

phenotypic traits. For example, CNNs can process images of seeds to detect morphological characteristics associated with high or low germination rates.

- ✓ **Seed Viability Assessment:** ML algorithms can assess seed viability by analyzing seed morphology, color and texture from digital images. Advanced models can differentiate between viable and non-viable seeds based on subtle differences in appearance that may not be visible to human inspectors.
- ✓ **Disease and Pest Detection:** Seed-borne diseases and pests can significantly affect seed quality. Machine learning techniques such as SVMs and deep learning models can identify diseased seeds from healthy ones by analyzing patterns in spectral data or seed images. These techniques can also detect fungal infections, bacterial contamination or insect infestations.
- ✓ **Seed Purity Analysis:** Seed purity is a measure of the proportion of desired seeds in a sample excluding foreign materials, weed seeds or other crop seeds. ML algorithms can automate the process of purity analysis by classifying seed types based on their physical or spectral characteristics.
- ✓ **Moisture Content Determination:** Moisture content is a key determinant of seed storability and germination potential. Machine learning models combined with near-infrared (NIR) spectroscopy can accurately predict moisture content levels in seeds without the need for destructive testing.

Key Machine Learning Models for Seed Quality Determination

Machine learning encompasses a wide range of models each with unique strengths and applications. Some of the most prominent models used in seed quality determination include:

- **Convolutional Neural Networks (CNNs):**

CNNs are a class of deep learning models specifically designed for image analysis. They consist of multiple layers that automatically learn to detect features in images such as edges, textures and shapes. CNNs have been extensively used for seed quality determination by analyzing seed images to detect defects, classify seed types and assess seed viability.

For example, CNN-based approach to classify soybean seeds into different quality categories based on their visual characteristics. The model achieved an accuracy of over 95%, demonstrating the potential of CNNs in automating seed quality assessment.

JUST AGRICULTURE
multidisciplinary e-Newsletter

- **Support Vector Machines (SVMs):**

SVMs are supervised learning models used for classification and regression tasks. They are particularly effective in high-dimensional spaces and can handle non-linear relationships between variables. In seed quality determination SVMs have been used to classify seeds based on spectral data, morphological traits and other features.

A research study utilized SVMs to classify wheat seeds into healthy and diseased categories based on hyperspectral imaging data. The model achieved high classification accuracy indicating its effectiveness in detecting seed-borne diseases.

- **Random Forests:**

Random forests are ensemble learning models that combine multiple decision trees to improve classification accuracy and reduce overfitting. They are robust to noise and can handle large datasets with numerous input variables. Random forests have been applied in seed quality determination for tasks such as purity analysis, moisture content prediction and germination rate estimation.

For instance, a study used a random forest model to predict the germination rates of maize seeds based on phenotypic traits and environmental factors. The model provided accurate predictions helping farmers make informed decisions about seed selection.

- **K-Nearest Neighbors (KNN):**

KNN is a simple, yet effective, machine learning algorithm used for classification and regression. It works by finding the 'k' nearest data points to a query point and making predictions based on their labels. In seed quality analysis, KNN has been used to classify seed types, detect defects and assess viability.

A study employed KNN to classify rice seeds into different quality categories based on their morphological characteristics. The model showed high accuracy, demonstrating the potential of KNN in automating seed quality assessment.

Advantages of Machine Learning in Seed Quality Determination

Machine learning offers several advantages over traditional methods for seed quality determination:

- **Accuracy and Precision:** ML models can achieve higher accuracy and precision in seed quality assessment by analyzing large datasets and identifying subtle patterns that are not detectable by human inspectors.

JUST AGRICULTURE
multidisciplinary e-Newsletter

- **Speed and Efficiency:** Machine learning algorithms can process vast amounts of data in a fraction of the time required for manual analysis, enabling real-time quality assessment.
- **Consistency:** Unlike human inspectors, ML models provide consistent results, reducing variability in quality assessment due to human error or subjective judgment.
- **Scalability:** ML-based systems can be easily scaled to handle large volumes of seeds, making them suitable for large-scale seed testing laboratories and seed production companies.
- **Cost-Effectiveness:** While the initial investment in machine learning technology may be high, the long-term cost savings due to reduced labor costs, faster processing times and improved accuracy make it a cost-effective solution.

Challenges and Limitations of Machine Learning in Seed Quality Determination

Despite its numerous advantages, machine learning also presents several challenges and limitations in the context of seed quality determination:

- **Data Requirements:** ML models require large amounts of high-quality data for training and validation. Collecting and labeling such data can be time-consuming and expensive.
- **Complexity:** Developing and implementing ML models can be complex, requiring specialized knowledge in data science, computer vision and seed physiology.
- **Model Generalization:** Machine learning models trained on specific datasets may not generalize well to new datasets with different environmental conditions or seed varieties.
- **Interpretability:** Some ML models especially deep learning models are considered 'black boxes,' making it difficult to interpret their decisions and understand the underlying patterns.
- **Infrastructure and Cost:** Implementing machine learning solutions requires significant investment in computational infrastructure, software, and expertise, which may not be feasible for all organizations, particularly small-scale seed producers.

Future Directions in Machine Learning for Seed Quality Assessment

As technology advances, several promising developments are expected in the field of machine learning for seed quality determination:

JUST AGRICULTURE
multidisciplinary e-Newsletter

- **Integration with IoT and Remote Sensing:** Combining ML with Internet of Things (IoT) devices and remote sensing technologies can provide real-time data for seed quality assessment, improving accuracy and efficiency.
- **Use of Transfer Learning:** Transfer learning, where a model trained on one task is adapted to perform another related task, can reduce the data requirements for training ML models, making them more accessible to smaller organizations.
- **Development of Explainable AI Models:** Researchers are working on developing explainable AI models that can provide insights into their decision-making processes, increasing trust and adoption of ML-based solutions.
- **Personalized Seed Quality Models:** Advances in ML could lead to the development of personalized seed quality models tailored to specific crops, regions, and farming practices, enhancing their applicability and effectiveness.
- **Increased Use of Unsupervised Learning:** Unsupervised learning methods, which do not require labeled data, could be used to identify novel patterns and correlations in seed quality data, further improving assessment methods.

Conclusion

Machine learning represents a significant advancement in the field of seed quality determination, offering numerous benefits over traditional methods, including improved accuracy, efficiency, and scalability. Despite the challenges associated with its adoption, the potential of ML to revolutionize seed quality assessment is undeniable. As research progresses and technology becomes more accessible, machine learning is poised to play an increasingly important role in ensuring the availability of high-quality seeds, contributing to global food security and sustainable agriculture.

References:

De Medeiros, A.D.; Pinheiro, D.T.; Xavier, W.A.; da Silva, L.J.; dos Dias, D.C.F. Quality classification of Jatropha curcas seeds using radiographic images and machine learning. Ind. *Crops Prod. 2020*, 146, 112162.

ElMasry, G.; Mandour, N.; Al-Rejaie, S.; Belin, E.; Rousseau, D. Recent Applications of Multispectral Imaging in Seed Phenotyping and Quality Monitoring—An Overview. *Sensors 2019*, 19, 1090.

Finch-Savage, W.E.E.; Bassel, G.W.W. Seed vigour and crop establishment: Extending performance beyond adaptation. *J. Exp. Bot*. 2016, 67, 567–591.

Medeiros AD, Silva LJD, Ribeiro JPO, Ferreira KC, Rosas JTF, Santos AA, Silva CBD. Machine Learning for Seed Quality Classification: An Advanced Approach Using Merger Data from FT-NIR Spectroscopy and X-ray Imaging. *Sensors* (Basel). 2020 Aug 3;20(15):4319. doi: 10.3390/s20154319.

Wakholi, C.; Kandpal, L.M.; Lee, H.; Bae, H.; Park, E.; Kim, M.S.; Mo, C.; Lee, W.H.H.; Cho, B.K.K. Rapid assessment of corn seed viability using short wave infrared line-scan hyperspectral imaging and chemometrics. *Sens. Actuators B Chem.* 2018, 255, 498–507.

Xia, Y.; Xu, Y.; Li, J.; Zhang, C.; Fan, S. Recent advances in emerging techniques for non-destructive detection of seed viability: A review. *Artif. Intell. Agric*. 2019, 1, 35–47.