EPJ Web of Conferences **344**, 01061 (2025) https://doi.org/10.1051/epjconf/202534401061
AIPTEC 2025

Development of a deep learning-based expert system for early detection of corn diseases using the TF-IDF and Multi-Layer Perceptron approaches

Hanifudin Sukri[1*], *Puteri* Nurul Ma'rifah[2], *Ach.* Dafid [3], *Mahmud*[1], *Elfon* Septiyananda Alim Cahyono[3]

[1]Departement of Information System, Universitas Trunodjoyo Madura, Bangkalan, Indonesia
[2]Departement of Electrical Engineering, Universitas Trunodjoyo Madura, Bangkalan, Indonesia
[3]Departement of Mechatronics Engineering, Universitas Trunodjoyo Madura, Bangkalan, Indonesia

Abstract. Corn is a vital agricultural commodity, yet early disease detection remains challenging due to the linguistic variability in describing symptoms. To address this, this study proposes a deep learning-based text classification model designed to diagnose corn diseases using descriptive symptom inputs. The research contribution is the development of a robust intelligent system capable of accurately interpreting natural language symptom descriptions to overcome the limitations of rigid rule-based diagnostics. The methodology employs Term Frequency-Inverse Document Frequency (TF-IDF) for feature extraction combined with a Multi-Layer Perceptron (MLP) architecture. To ensure model robustness and generalization, data augmentation and the Synthetic Minority Over-sampling Technique (SMOTE) are applied to balance and expand the training dataset. The evaluation results demonstrate that the proposed TF-IDF and MLP model achieved excellent performance with an accuracy of 99.82%. The confusion matrix analysis indicates that precision, recall, and F1-score values were all equal to 1.00 across disease categories. Furthermore, the trained model was successfully converted into TensorFlow Lite (tflite) format for mobile deployment. Finally, the system was integrated into an Android-based mobile application named JagungKu to provide real-time diagnostic results. In conclusion, this research advances intelligent text-based disease detection systems and demonstrates the potential of deep learning in supporting sustainable precision agriculture.

1 Introduction

1.1 Background

Corn is one of the pillars of national food security in Indonesia, whose role goes beyond being merely a staple food source [1-3]. Its multifunctional role makes it a key commodity that not only provides food for the community, but also plays a vital and irreplaceable role in supporting the livestock sector as the main raw material for the feed industry [4-5]. The tangible contribution of this commodity to the national economy is reflected in data from the Badan Pusat Statistik (BPS), which shows a positive production trend, from 14.7 million tons in 2023 to around 15 million tons in 2024 [6]. This increase is not merely a statistic, but concrete evidence of corn's significant contribution to supporting the economy, from the farmer level to the downstream industry. With its strategic position in the food supply chain and industry, efforts to maintain and increase corn availability and productivity are key factors in determining national food stability and the sustainability of various agriculture-based industries in Indonesia.

However, corn production in Indonesia still faces various obstacles, one of which is plant disease that can significantly reduce crop yields [7-8]. Diseases such as gray spot, smut, and leaf rust often attack corn plants, causing damage to the leaves, stems, and cobs [9-10]. Production losses due to this disease not only have a direct impact on national food security, but also erode farmers' incomes and have the potential to trigger price instability in the market [11-12]. As technology advances, artificial intelligence-based approaches are beginning to be utilized in agriculture, whether for disease detection, harvest management, or crop condition prediction [13-15]. In this context, technology can be utilized to provide smart solutions for early detection of corn diseases in order to minimize potential production losses.

Previous research has developed a rule-based expert system that integrates the forward chaining method for the inference process and certainty factors to accommodate uncertainty in symptoms when diagnosing corn plant diseases. In terms of performance, the system has been proven capable of providing fairly reliable diagnostic results with an accuracy rate of 84% [16]. However, despite its accuracy, this system has fundamental limitations, particularly in terms of its flexibility in accepting input data. Rigid rule-based models can only operate effectively if the symptoms entered by the user exactly match the pre-programmed structured list of symptoms [17-18]. As a result, the

* Corresponding author: hanifudinsukri@trunojoyo.ac.id

system becomes highly vulnerable to linguistic variations. The system is unable to interpret or recognize new symptom descriptions that are not listed in the rule base, let alone understand free and natural sentence-based input from users. This rigidity is the main bottleneck, causing the system's ability to adapt to the diversity of delivery methods and symptom variations in the field to be very low. Thus, although accurate for the appropriate data, this system has a large gap in terms of user experience and application in dynamic real-world conditions.

To overcome these limitations, this study proposes the development of an innovative corn disease classification model based on text description by combining the Term Frequency–Inverse Document Frequency (TF-IDF) word weighting technique and the Multi-Layer Perceptron (MLP) artificial neural network algorithm. In the proposed architecture, TF-IDF acts as an intelligent feature extractor by identifying and weighting the most significant and relevant keywords from the symptom description text, while reducing the influence of common and uninformative words [19-21]. Furthermore, MLP, with its ability to model complex non-linear relationships, functions as a classifier that processes the best features of TF-IDF to map a set of textual symptoms into the most likely disease categories [22-23]. This research is expected to significantly improve the accuracy and flexibility of early detection of corn diseases, as well as provide a strong computational foundation as a strategic first step toward the application of artificial intelligence based on natural language processing (NLP) in Indonesia's agricultural sector.

1.2 Research objective

This study aims to develop a corn disease classification model based on text descriptions as an extension of the previous rule-based expert system, with a focus on overcoming its limitations in accepting static inputs and its inability to adapt to variations in natural language used by users. Through a deep learning approach, this study utilizes the TF-IDF technique, which is then processed by a Multi-Layer Perceptron (MLP) to identify complex patterns and determine disease types. To improve data quality, this study also applies augmentation techniques and the Synthetic Minority Over-sampling Technique (SMOTE) to balance the uneven distribution of disease classes. The final result is not only a classification model that is more flexible and accurate than conventional systems, but also an Android-based mobile application that can be easily implemented by farmers and agricultural extension workers for early detection of corn diseases using natural text descriptions.

1.3 Contribution

This research contributes both scientifically and practically to the development of an artificial intelligence-based corn disease detection system. Methodologically, this research introduces an innovative approach through a combination of TF-IDF and Multi-Layer Perceptron (MLP) techniques that can identify diseases from text descriptions, overcoming the limitations of static rule-based expert systems. The practical contribution is realized through the development of a dataset enriched with augmentation and SMOTE techniques, as well as the implementation of the model into a user-friendly Android mobile application for farmers and agricultural extension workers. Thus, this research not only enriches the application of NLP and Deep Learning in agriculture but also serves as a strategic step toward the digital transformation of Indonesia's agricultural sector through widely accessible smart technology.

2 Literature review

2.1 Text-based corn disease detection

The text-based plant disease detection approach applies Natural Language Processing (NLP) that focuses on processing symptom descriptions provided by users in the form of natural sentences. In the context of corn plants, symptom descriptions such as changes in leaf color, abnormal growth, or damaged stems and cobs can be analyzed to recognize linguistic patterns that indicate specific types of diseases. By converting text into numerical data that can be studied by artificial intelligence models, the system can understand the context and meaning of the natural language used by farmers or field experts without having to follow a standard input format. This approach makes the diagnosis process more adaptive and intuitive, as the system is able to adapt to variations in user language and expand its knowledge as more data is added, thereby supporting the development of a smarter agricultural disease detection system that is more responsive to real conditions in the field.

2.2 Transition from expert system to machine learning

The development of artificial intelligence technology has driven a significant paradigm shift, from rigid rule-based expert systems to more adaptive machine learning-based systems [24-25]. Rule-based systems, which are deterministic and rely on a series of if–else logic constructed manually by experts, have a fundamental weakness in adapting to new variations and patterns outside their initial knowledge base [26]. In contrast, the machine learning approach offers a solution with its ability to automatically learn complex and non-linear patterns of relationships between features directly from the data, thereby eliminating the dependence on explicit rules that must be defined in advance [27]. In the context of this study, the paradigm shift was realized by utilizing a collection of symptom data that previously formed the foundation of the expert system. The data was reprocessed through text-to-vector transformation using the Term Frequency–Inverse Document Frequency (TF-IDF) method to capture the statistical significance of each word, which then became input for the Multi-Layer Perceptron (MLP) algorithm to learn hidden diagnostic patterns. This hybrid TF-IDF and MLP approach has proven to be robust and successfully applied in various domains such as medical document classification, educational sentiment analysis, and environmental quality monitoring, where deep learning models consistently show significant improvements in classification accuracy and system generalization capabilities on previously unseen data. Therefore, a similar transition in corn disease diagnosis is expected to not only optimize accuracy, but also create a system capable of evolving and adapting dynamically by learning the characteristics of descriptive symptom data that continue to grow and change over time.

2.3 TF-IDF in text classification

Term Frequency-Inverse Document Frequency (TF-IDF) is one of the most commonly used statistical methods in Information Retrieval and Text Mining to measure the importance of a term in a document relative to a collection of documents (corpus). The main strength of TF-IDF lies in its ability to quantify the informative value of a word by considering two aspects: its prevalence in the local document and its scarcity in the global corpus. Mathematically, this method is the product of two components. First, Term Frequency (TF), which counts the frequency of a word's appearance in a particular document. A high TF value indicates that the word is dominant and may be central to describing the content of the document. Second, Inverse Document Frequency (IDF), which measures how rare or common a word is in all documents in the corpus. Words that appear in almost all documents (such as conjunctions or common words) will have a low IDF value, as they are considered to have no distinguishing power. By multiplying TF and IDF, TF-IDF produces a numerical weight that effectively filters out common “noisy” words and simultaneously increases the weight of specific and informative keywords [28-29]. In the context of this research, TF-IDF serves as a critical transformation tool for converting unstructured descriptive text data on disease symptoms into structured numerical vector representations. These vectors form the feature space in which the classification model can operate. The advantages of TF-IDF lie in its simplicity, computational efficiency, and proven ability to capture relevant distinguishing words, making it a solid choice as a foundation for text feature extraction.

2.4 Multi-Layer Perceptron (MLP) in classification

Multi-Layer Perceptron (MLP) is a fundamental architecture in artificial neural networks designed to handle data that cannot be separated linearly [30]. The MLP structure consists of three main components: an input layer that receives data features, one or more hidden layers that perform complex computations, and an output layer that generates the final prediction [31-32]. For illustration, see Fig. 1. below.

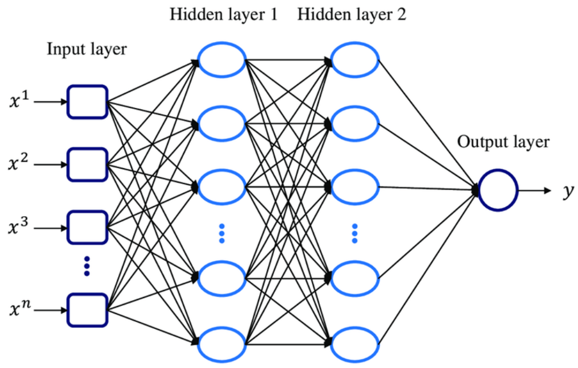

Fig. 1. Simple MLP architecture [33].

The connections between neurons in these layers have trainable weights, whose values are adjusted iteratively during training using the backpropagation algorithm to minimize the loss function. The capacity of MLP to model non-linear relationships comes from the application of activation functions in each hidden layer, such as Rectified Linear Unit (ReLU), which introduces non-linearity into the network. For multi-class classification tasks such as disease diagnosis, the output layer typically uses the Softmax activation function, which converts values into a probability distribution for each disease class.

In the context of text classification, after text data is converted into numerical representations, MLP acts as a classifier that learns the mapping from these input features to target categories. The main advantage of MLP lies in its ability to approximate highly complex non-linear functions between input features. This makes it particularly well-suited for handling the variations and nuances in the description of plant disease symptoms, where the relationship between certain words and a disease may not be direct or may depend on context. Its ability to learn hierarchically through hidden layers allows the MLP to capture abstract patterns from feature extraction data, resulting in a model that is more robust and accurate than simple linear models.

2.5 Research gap and novelty

This research is a direct development of previous research that has implemented a rule-based expert system with forward chaining and certainty factor methods for diagnosing corn plant diseases. The system achieved an accuracy rate of 84% [16]. However, it still has limitations because it can only process symptom inputs in the form of static options and cannot recognize symptom descriptions written naturally by users. This weakness is the starting point for current research, which focuses on transforming the rule-based approach into a deep learning approach through text representation.

Several recent studies show that combining text representation techniques and deep learning-based classification algorithms has great potential for improving the accuracy of text-based diagnostic systems. Luo and Lu (2024) used a combination of TF-IDF and the Naïve Bayes algorithm to improve natural language classification accuracy through more contextual word weighting [34]. Liang and Niu (2022) refined TF-IDF by considering word position and lexical category, resulting in improved accuracy in context-based text classification [35]. Sirisha et al. (2024) introduced Dynamic Multi-Layer Perceptron (DMLP) for fetal health classification using cardiotocography data, and showed that the addition of adaptive hidden layers could improve accuracy to 97% [36]. Research by Han et al. (2025) combined MLP with Graph Neural Network (GNN) for metastatic prostate cancer detection, and successfully improved accuracy to 88% through complex feature pattern learning [37]. Meanwhile, Çakan et al. (2024) demonstrated the efficiency of the MLP decomposition architecture in processing IoT data, which is capable of achieving 90% of the system's maximum throughput [38].

From these various studies, it can be concluded that TF-IDF is suitable for extracting key features from text with relevant word weights, while MLP excels at learning non-linear relationships between features to produce more accurate classifications. However, there has been no research that applies a combination of TF-IDF and MLP in an integrated manner to diagnose plant diseases, especially in corn, based on text descriptions of symptoms. Thus, this study presents a new approach by utilizing TF-IDF as a feature extractor and MLP as a classifier on text-based corn disease symptom data. The resulting model is expected to be able to recognize descriptive linguistic patterns of symptoms flexibly and accurately, without relying on static rules as in previous expert systems. This approach not only improves disease detection accuracy but also represents an innovative step toward the application of Natural Language Processing (NLP) and Deep Learning in the Indonesian agricultural sector.

3 Methodology

The methodology of this study consists of several main stages, namely data preparation, preprocessing, feature extraction, model training, and evaluation, as shown in Fig. 1. The process begins with transforming expert-based symptom and disease data into textual form, followed by text augmentation techniques such as synonym replacement, random sentence shuffle, translation, and backtranslation to enrich data diversity. During preprocessing, the text undergoes translation, case folding, cleaning, tokenization, stemming, and stopword removal to standardize input. The resulting clean text is converted into numerical vectors using TF-IDF, and class imbalance is handled using SMOTE. The processed data is then trained using a Multi-Layer Perceptron (MLP) model, and the system's performance is evaluated through accuracy, precision, recall, and F1-score metrics derived from the confusion matrix.

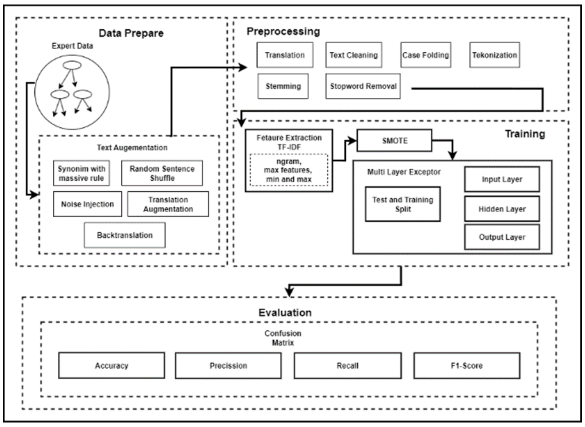

Fig. 1. The methodology of this study.

3.1 Data collection

The data used in this study came from the knowledge base of the corn disease diagnosis expert system in previous studies. The initial data source consisted of unique symptom codes (G1-G24) describing the physical condition of the plants and seven disease categories (P01-P07) that formed symptom-disease relationships through expert rules. This initial dataset then underwent a comprehensive validation and expansion process by phytopathologists to ensure its suitability for field conditions.

The validation process included re-examining symptom descriptions in accordance with standard agronomic terminology, adjusting overlapping disease categories, and adding new symptoms related to fungal and bacterial infections that are common in humid tropical conditions. The result of this process is a more complete, consistent, and representative dataset.

This validated final dataset will form the basis for the preprocessing and training of Deep Learning models using a combination of TF-IDF and MLP in the next stage of research.

3.2 Data augmentation

The data augmentation stage was carried out to expand the variety of descriptions of corn disease symptoms, so that the trained model did not only rely on the original text patterns from the expert system results, but was also able to recognize various natural language variations [39] [40]. This augmentation process is carried out through an automated Natural Language Processing (NLP)-based pipeline that combines several generative techniques. Each piece of text data transformed from the expert system is expanded with new variations using agriculture-specific synonym replacement rules [41] [42], For example, as shown in Table 1 below.

Table 1. Specifif synonym replacement.

At first	*Synonim*
Yellow stripes on leaves	yellow lines on the leaves
The leaves are whitish yellow	yellowish-white leaves
Small cobs	small corn cob

In addition, augmentation also applies random sentence shuffle, which is the randomization of sentence order in a description to produce different narrative structures without changing the core meaning of the symptom. For example, in Table 2 below.

Table 2. Random sentence shuffle.

At First	*Random Shufle*
Corn plants show symptoms of yellow stripes on leaves, indicating leaf blight. The leaves appear stiff and growth is stunted.	The leaves appear stiff and growth is stunted. Corn plants show symptoms of yellow stripes on leaves, which indicates leaf blight.
Leaf spot symptoms on corn are characterized by round to oval patches and yellow spots on leaves.	Yellow spots on leaves and round to oval patches are signs of leaf spot on corn.

To enrich linguistic diversity, this study uses back-translation and translation augmentation techniques, which involve translating back and forth between languages to create natural sentence variations with different word equivalents. For example, in Table 3 below.

Table 3. Back-translation and translation augmentation.

Id	En	Fr	Id
Tanaman jagung menunjukkan gejala daun bergaris kuning yang mengindikasikan Bulai.	Corn plants show yellow striped leaves indicating downy mildew.	Les plants de maïs présentent des feuilles rayées jaunes indiquant le mildiou.	Tanaman jagung tampak memiliki daun dengan garis kuning yang menandakan penyakit Bulai.

Each augmented text is then combined and selected to eliminate duplication, resulting in thousands of unique variations of symptom combinations with diseases.

3.3 Text preprocessing

The text preprocessing stage is carried out to ensure that all symptom texts have a uniform format and are ready for use. Each augmented text is processed through several stages that are integrated sequentially, as follows

3.3.1 Translation

Here, we translate the entire text into Indonesian to standardize the language. As shown in Table 4.

Table 4. Translation.

Before	After
Corn plants show yellow striped leaves indicating downy mildew.	Tanaman jagung menunjukkan daun bergaris kuning yang mengindikasikan bulai.

3.3.2 Case folding

By changing all letters to lowercase. The goal is to standardize the text format. As shown in Table 5**.**

Table 5. Case Folding.

Before	After
Tanaman Jagung Menunjukkan Daun Bergaris Kuning Yang Mengindikasikan Bulai.	tanaman jagung menunjukkan daun bergaris kuning yang mengindikasikan bulai

3.3.3 Text cleaning

This stage serves to remove excessive characters, non-alphabetic characters such as numbers, punctuation marks, special symbols, double spaces, and residual symbols from the previous process. As shown in Table 6.

Table 6. Text cleaning.

Before	After
tanaman jagung menunjukkan daun--- bergaris kuning !!	tanaman jagung menunjukkan daun bergaris kuning

3.3.4 . Tokenization

Breaking text into individual words (tokens). As shown in Table 7**.**

Table 7. Tokenization.

Before	After
tanaman jagung menunjukkan daun bergaris kuning yang mengindikasikan bulai	[‘tanaman’, ‘jagung’, ‘menunjukkan’, ‘daun’, ‘bergaris’, ‘kuning’, ‘yang’, ‘mengindikasikan’, ‘bulai’]

3.3.5 Stemming

Stemming converts each token to its base form by removing affixes or suffixes. As shown in Table 8**.**

Table 8. Stemming.

Before	After
[‘tanaman’, ‘jagung’, ‘menunjukkan’, ‘daun’, ‘bergaris’, ‘kuning’, ‘mengindikasikan’, ‘bulai’]	[‘tanam’, ‘jagung’, ‘tunjuk’, ‘daun’, ‘garis’, ‘kuning’, ‘indikasi’, ‘bulai’]

3.3.6 Stopword removal

Removing stopwords or common words that do not have high informational value, such as “which,” “and,” “with,” “on,” and others, so that the model only focuses on important words. As shown in Table 9**.**

Table 9. Stopword removal.

Before	After
[‘tanam’, ‘jagung’, ‘tunjuk’, ‘daun’, ‘garis’, ‘kuning’, ‘indikasi’, ‘bulai’]	[‘tanam’, ‘jagung’, ‘daun’, ‘garis’, ‘kuning’, ‘bulai’]

After all preprocessing steps have been applied, the descriptive text on corn diseases is transformed into a more structured form and is ready for the next process.

3.4 Feature extraction using TF-IDF

The feature extraction stage is carried out using the Term Frequency–Inverse Document Frequency (TF-IDF) method to convert the preprocessed text into a numerical representation that can be understood by machine learning models. TF-IDF works by assigning a weight to each word based on its frequency of occurrence in a single document (term frequency) and its scarcity across the entire document collection (inverse document frequency) [28-29]. The more frequently a word appears in a particular document but rarely appears in other documents, the greater its weight as a representative word. In the context of this study, each description of corn disease symptoms was converted into a high-dimensional numerical vector representing the level of importance of each word to a particular disease category. The results of this TF-IDF representation are then used as input in the Multi-Layer Perceptron (MLP) model for automatic disease classification, because this vector format allows the model to recognize patterns of relationships between words that are relevant in detecting types of corn plant diseases.

3.5 SMOTE

The augmented dataset has several disease categories with far fewer samples than other classes. To address this issue, this study uses the Synthetic Minority Over-sampling Technique (SMOTE), which balances the proportion of data between classes by adding synthetic samples to minority classes.

Unlike oversampling techniques that duplicate data, SMOTE works by generating new synthetic data between existing minority data points [43]. This approach not only increases the amount of minority class data, but also helps expand the feature space so that the model can better learn data distribution patterns and avoid overfitting. For illustration, see the Fig. 2 below.

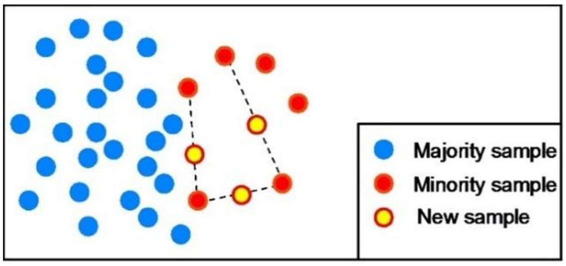

Fig. 2. How SMOTE works [43].

3.6 Multi-Layer Perceptron (MLP)

The main classification approach used in this study is Multi-Layer Perceptron (MLP), which is a type of artificial neural network consisting of several interconnected layers, including an input layer, several hidden layers, and an output layer. The input layer receives the numerical vector resulting from TF-IDF extraction, then processes it through the hidden layer using a non-linear activation function such as ReLU (Rectified Linear Unit) to learn the complex relationships between features [30]. Next, the output layer uses the Softmax activation function to generate probabilities for each disease class, allowing the model to determine the disease category with the highest probability value. The learning process is carried out using the backpropagation algorithm, where the weights between neurons are adjusted iteratively using the categorical cross-entropy loss function and optimized with the Adam Optimizer algorithm. This architecture allows the MLP to recognize non-linear and variable symptom patterns from descriptive text, thereby improving the accuracy of corn disease detection compared to static rule-based methods [44].

3.7 Evaluation

Model evaluation was conducted to measure the extent to which the developed classification system was able to recognize corn plant diseases based on text descriptions with a good level of accuracy [45]. This study uses four main metrics, namely accuracy, precision, recall, and F1-score, each of which is calculated based on the confusion matrix results. The confusion matrix is a table that describes the comparison between the model's prediction results and the actual values (ground truth), which consists of four main components [46-47]. True Positive (TP) is the number of positive data that are correctly predicted, True Negative (TN) is negative data that are correctly identified, False Positive (FP) is negative data that are incorrectly predicted as positive, and False Negative (FN) is positive data that are not detected by the model [48].

Accuracy is defined as the proportion of correctly classified samples to the total number of predictions. It provides a general overview of the model’s performance across all sentiment categories [49]. However, in the context of imbalanced datasets where certain classes may dominate accuracy alone may not fully reflect the model's reliability. The formula used to compute accuracy is presented below.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Precision measures the proportion of correctly predicted positive instances out of all instances predicted as positive [50]. This metric is particularly important in contexts where false positives may carry significant consequences. In this study, precision was calculated individually for each sentiment class (positive, neutral, negative) to evaluate how accurately the model identified true sentiments without overgeneralizing. The formula for precision is shown below.

$$Precission = \frac{TP}{TP + FP} \tag{2}$$

Recall, also known as sensitivity, represents the model's capability to correctly identify all relevant instances of a given class [51]. A high recall indicates that the model successfully captures most of the actual sentiment labels. This metric is particularly valuable for ensuring that minority sentiment classes are not overlooked. The formula to compute recall is provided below.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

The F1-score, calculated as the harmonic mean of precision and recall, providing as a balanced indicator of model performance [52]. It is especially useful in situation involving class imbalance, as it accounts for both false positives and false negatives. It is particularly useful in cases of class imbalance, where relying on a single metric may give a skewed impression of model performance. A higher F1-score indicates more robust overall classification performance. The formula for the F1-score is outlined below.

$$F1 = 2 \times \frac{Precission \times Recall}{Precisison + Recall} \tag{4}$$

3.8 Modeling

The modeling stage is the process of integrating all the main components of the system [53], which consists of a complete pipeline ranging from text preprocessing, feature extraction using TF-IDF, model training with Multi-Layer Perceptron (MLP), to producing output in the form of disease classification results. After the model achieves optimal performance based on the evaluation results, the trained model is then exported and converted into TensorFlow Lite (.tflite) format. This format was chosen because it is lightweight and efficient, making it ideal for implementation on Android-based mobile devices. Thus, the preprocessing → TF-IDF → MLP → classification results (tflite) pipeline allows the system to work end-to-end in mobile applications, where users simply enter a text description of their symptoms, and the system automatically provides a diagnosis of corn disease.

3.9 Mobile application implementation

The system is implemented in the form of an Android-based mobile application that functions as a user interface to detect corn plant diseases based on text descriptions of symptoms. The integration process is carried out by placing the .tflite model in the application's asset directory, then accessed through the TensorFlow Lite Interpreter on the frontend. When the user enters a symptom description via a text field, the application performs simple preprocessing locally, extracts features using a precompiled TF-IDF representation, and sends the resulting vector to the model for prediction. The model output is the probability of each disease class, and the result with the highest value is displayed as the primary diagnosis to the user.

4 Result and discussion

4.1 Data collection and augmentation

The initial data for this study was obtained from the results of a previous expert system containing pairs of 24 symptoms with 7 diseases, each of which was associated with one or more disease categories. This data was then expanded with new disease information with symptom data and disease data as shown in the Table 10 and Table 11 below.

Table 10. Symptom data.

Symptoms	
Yellow stripes on leaves	Brown color on the lower stem nodes
White flour covered line	Stems wet, soft, and Ringed
The leaves are whitish yellow	Smells bad
Stiff leaves	In the middle of the brown spots
Dwarf Plants / Short stems	The seeds contain white to gray brown mushrooms
Cob formation is inhibited	Rot is pink or reddish to brown
Small cobs	Leaves curl or roll inward
Round to oval patches	Presence of white powdery layer on leaf surface
Yellow spots on leaves	Leaves with necrotic edges
Plants Light brown to dark	Dark green streaks before yellowing
Wet plant	Moldy growth on cob surface
Small rusty red spots/stains	Seedlings fail to emerge or die after sprouting
Dotted lines all over the leaf surface	Brown lesions on midrib
Spots are gray / whitish	Roots rot and turn black
Spots on the fronds	Wilting during the day but recovers at night
Salmon color blotches	Leaf veins appear chlorotic (yellowed)
Widespread patches of gray or white color	Cobs produce lightweight or empty grains
Plants break suddenly	

Table 11. Disease data.

Disease	
Bulai	Leaf Spot
Leaf Rust	Dwarf Virus
Leaf Rot	Stem Rot
Cob Rot	Powdery Mildew
Bacterial Leaf Blight	Root Rot
Maize Streak Virus	Seedling Blight

The data was then expanded using an NLP-based augmentation process and balanced between disease classes using the Synthetic Minority Over-sampling Technique (SMOTE) method. The results of these two stages produced a total of 7,572 descriptive data points that were evenly distributed across each disease class, as shown in **Fig. 3**, where each disease category had a balanced number of samples.

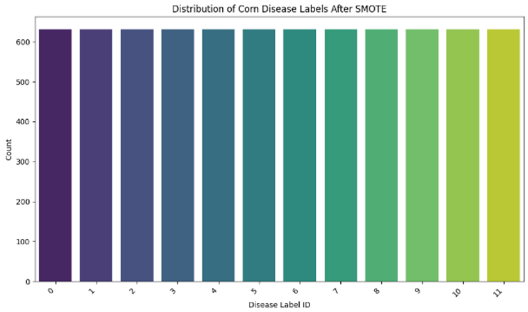

Fig. 3. Distribution of Corn Disease Labels After Augmentation and SMOTE.

4.2 Evaluation of MLP model

Model performance evaluation was conducted using four main metrics, namely accuracy, precision, recall, and F1-score, each of which was measured based on the confusion matrix results for all corn disease classes. The test results show that the combination of TF-IDF, SMOTE, and MLP was able to achieve an accuracy level of 99.82%, with precision, recall, and F1-score values consistently close to perfect across all disease classes. Based on Fig. 4, each disease category was correctly classified without any significant prediction errors.

	precision	recall	f1-score	support
P01	1.00	1.00	1.00	158
P02	1.00	1.00	1.00	120
P03	1.00	1.00	1.00	69
P04	1.00	1.00	1.00	138
P05	0.98	1.00	0.99	64
P06	0.99	1.00	0.99	97
P07	1.00	1.00	1.00	85
P08	1.00	1.00	1.00	81
P09	1.00	0.99	0.99	90
P10	1.00	0.99	0.99	81
P11	1.00	1.00	1.00	84
P12	1.00	1.00	1.00	57
accuracy			1.00	1124
macro avg	1.00	1.00	1.00	1124
weighted avg	1.00	1.00	1.00	1124

Fig. 4. Precision recall F1-score and support metric results.

The distribution of values in the confusion matrix also shows a full diagonal, indicating no classification shift between classes. This shows that the model is able to learn linguistic features well from the augmented and SMOTE-balanced data. The macro average and weighted average values, both of which reached 1.00, reinforce the evidence that the model has a very high generalization ability, even in classes with fewer original data. With these results, it can be concluded that the MLP architecture with TF-IDF vector input is a good approach for corn disease classification based on text descriptions.

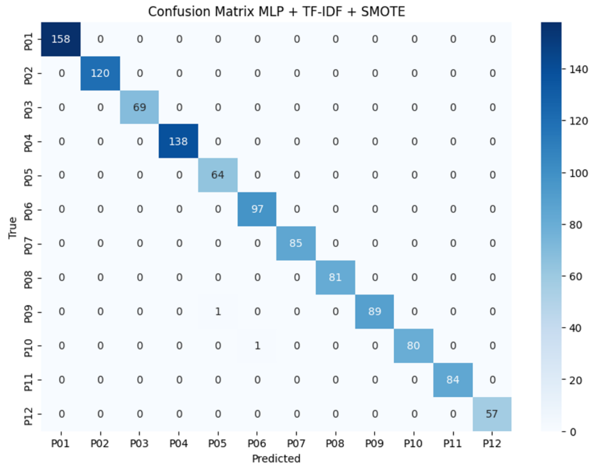

Fig. 5. Confusion matrix and result matirx evaluation.

4.3 Mobile application implementation

The corn disease classification model that has been converted to TensorFlow Lite is implemented in a mobile application called JagungKu. This application is designed to help farmers analyze plant symptoms quickly and accurately by simply entering a description of the symptoms in text form. The system architecture in this application runs offline, where the tflite model is called through the TensorFlow Lite Interpreter on the device side.

The JagungKu application has three main screens as shown in Fig. 6.The first screen is the home page, which contains an introduction to the application. The second screen is the symptom input page. And the last screen is the detection results page, which displays the identified disease type and percentage. The application interface is designed with a user-friendly approach that is convenient to use in the field and remains informative for farmers and agricultural extension workers.

4.4 Main findings of the present study

The primary finding of this research is that the integration of TF-IDF feature extraction with a Multi-Layer Perceptron (MLP) architecture is highly effective for diagnosing corn diseases based on textual symptom descriptions. The model achieved a near-perfect accuracy of 99.82%, with precision, recall, and F1-score values reaching 1.00 across all disease classes. This performance was largely driven by the application of SMOTE, which successfully equalized the distribution of minority classes, preventing the classifier from biasing towards the majority class.

4.5 Comparison with other studies

Compared to the previous rule-based expert system which achieved 84% accuracy, the proposed MLP model offers a substantial improvement. Rule-based systems struggle with linguistic variability because they rely on exact string matching. In contrast, the TF-IDF and MLP approach captures the semantic weight of keywords, allowing the system to process natural language variations more effectively. This aligns with recent studies suggesting that deep learning models, such as MLP, excel in learning non-linear relationships in complex datasets compared to linear or rule-based models.

4.6 Implication and explanation of findings

The high accuracy implies that textual data, when properly preprocessed and vectorized, contains sufficient information entropy to distinguish between specific plant pathologies. The success of the model can be explained by the MLP's hidden layers, which learn hierarchical representations of the symptom features extracted by TF-IDF. Practically, this finding supports the feasibility of deploying lightweight text-based diagnostic tools on mobile devices, as demonstrated by the *JagungKu* application, enabling farmers to obtain real-time decision support without the need for expensive imaging sensors.

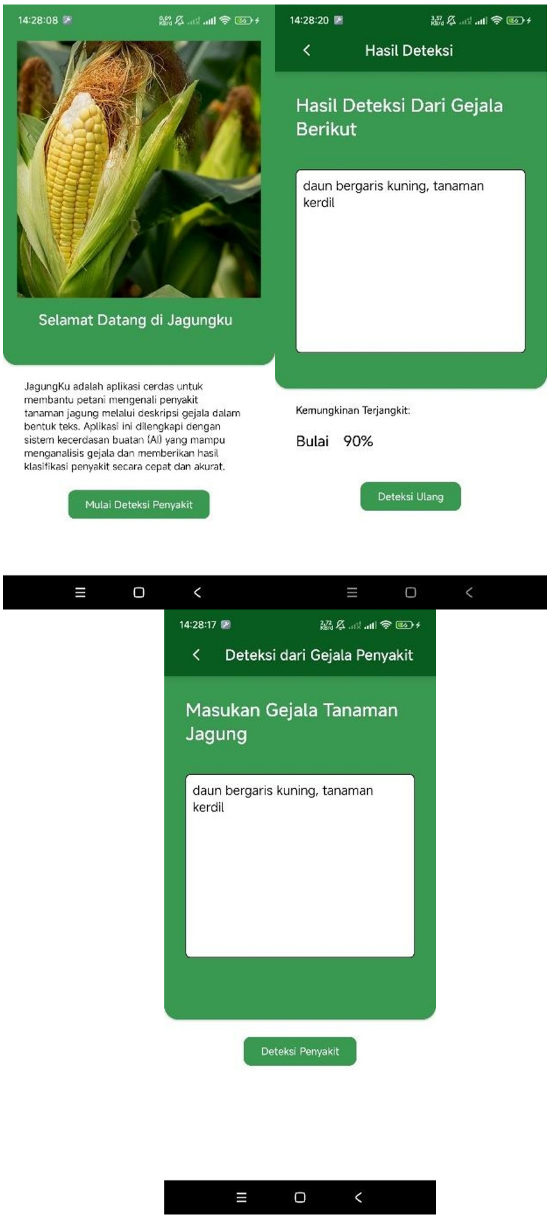

Fig. 6. JagungKu Mobile Application Implementation (a) Home Page, (b) Symptom Input, (c) Detection Results.

4.7 Strengths and limitations

The main strength of this study is the end-to-end implementation of a lightweight, offline-capable mobile diagnosis system that handles natural language, addressing the rigidity of traditional expert systems. However, the study has limitations. First, the model relies solely on textual input; thus, diagnostic accuracy is dependent on the user's ability to describe symptoms correctly. Second, while the dataset covers major corn diseases, it may not yet generalize to rare pathogens outside the current training corpus. Future research should address these by integrating multi-modal data (combining text and leaf imagery) to enhance robustness in diverse field conditions.

5 Conclusion and suggestions

This study successfully developed a text-based corn disease classification system using a combination of TF-IDF, SMOTE, and Multi-Layer Perceptron (MLP), achieving an exceptional accuracy of 99.82% with perfect precision, recall, and F1-score values. The model effectively transformed descriptive textual inputs into numerical representations, enabling accurate and adaptive disease detection based on linguistic patterns. Furthermore, the trained model was optimized and implemented into an Android-based application named JagungKu, allowing users to diagnose corn diseases quickly and efficiently through natural language input. The findings demonstrate that text-based analysis can serve as a reliable alternative for plant disease detection, especially in field conditions where visual data may be limited. For future research, the system can be expanded through integration with IoT-based monitoring or conversational AI, enabling real-time and interactive disease diagnosis that supports smart and sustainable agricultural practices.

References

1. Admindiskatan. Wahyu: Jagung untuk Ketahanan Pangan Nasional. diskatan.kuningankab.go.id. (2025). https://diskatan.kuningankab.go.id/wahyu-jagung-untuk-ketahanan-pangan-nasional/
2. W. Widhiantini, N. M. C. Sukendar, A. A. G. Purantara, Design of the Bali Province Food Security Action Plan towards Food Independence. J Environ Manag Tour. **15**, 1, 30 (2024). https://doi.org/10.14505/jemt.v15.1(73).03
3. A. A. Dzulhidany, M. S. A. Rahman, Cultivating Food Sovereignty in the Time of the Pandemic: An Analysis of Jokowi's Agricultural Policy. KnE Soc Sci. (2022). https://doi.org/10.18502/kss.v7i4.10523
4. A. M. Shah, H. Zhang, M. Shahid et al., The Vital Roles of Agricultural Crop Residues and Agro-Industrial By-Products to Support Sustainable Livestock Productivity in Subtropical Regions. Animals. **15**, 8, 1184 (2025). https://doi.org/10.3390/ani15081184
5. H. Enawgaw, T. Tesfaye, K. T. Yilma, D. Y. Limeneh, Multiple Utilization Ways of Corn By-Products for Biomaterial Production with Bio-Refinery Concept; a Review. Mater Circ Econ. **5**, 1, 7 (2023). https://doi.org/10.1007/s42824-023-00078-6
6. BPS Indonesia, Luas Panen, Produksi, dan Produktivitas Jagung Menurut Provinsi 2023–2024. bps.go.id. (2025). https://www.bps.go.id/id/statistics-

table/2/MjIwNCMy/luas-panen–produksi–dan-produktivitas-jagung-menurut-provinsi.html

7. N. I. Nadi, A. Halid, R. Mustafa, Analisis Risiko Produksi dan Pendapatan Usahatani Jagung di Desa Dulamayo Utara Kecamatan Telaga Biru Kabupaten Gorontalo. JIA J Agribisnis dan Ilmu Sos Ekon Pertan. **9**, 5, 500–510 (2024). https://doi.org/10.37149/jia.v9i5.1472
8. S. P. Sari, I. Suliansyah, N. Nelly, H. Hamid, I. Dwipa, Corn Pests and Evaluation of the Implementation of Integrated Pest Management in West Sumatra Indonesia. Int J Adv Sci Eng Inf Technol. **13**, 1, 91–96 (2023). https://doi.org/10.18517/ijaseit.13.1.16988
9. H. Wang, M. He, M. Zhu, G. Liu, Wcg-Vmamba: A Multi-Modal Classification Model for Corn Disease. Comput Electron Agric. **230**, 109835 (2024). https://doi.org/10.2139/ssrn.5016936
10. M. H. Lubis, N. Purnomo II, Identifikasi Penyakit Tanaman Jagung Dengan Metode Certainty Factor. J Sci Soc Res. (2024). https://doi.org/10.54314/jssr.v7i3.2080
11. N. S. Amin, A. Rauf, Y. Saleh, Diversifikasi Berbagai Tanaman Sela pada Budidaya Jagung di Kabupaten Bone Bolango. AGRINESIA J Ilm Agribisnis. 31–37 (2024). https://doi.org/10.37046/agr.v0i0.29416
12. P. Khatri, P. Kumar, K. S. Shakya, M. C. Kirlas, K. K. Tiwari, Understanding the Intertwined Nature of Rising Multiple Risks in Modern Agriculture and Food System. Environ Dev Sustain. **26**, 9, 24107–24150 (2023). https://doi.org/10.1007/s10668-023-03638-7
13. R. Rani, J. Sahoo, S. Bellamkonda, S. Kumar, S. K. Pippal, Role of Artificial Intelligence in Agriculture: An Analysis and Advancements with Focus on Plant Diseases. IEEE Access. **11**, 137999–138019 (2023). https://doi.org/10.1109/ACCESS.2023.3339375
14. İ. Yağ, A. Altan, Artificial Intelligence-Based Robust Hybrid Algorithm Design and Implementation for Real-Time Detection of Plant Diseases in Agricultural Environments. Biology. **11**, 12, 1732 (2022). https://doi.org/10.3390/biology11121732
15. M. R. Tammina, K. Sumana, P. P. Singh, T. R. V. Lakshmi, S. D. Pande, Prediction of Plant Disease Using Artificial Intelligence. , 25–48 (2024). https://doi.org/10.1007/978-981-99-9621-6_2
16. H. Sukri, F. Adiputra, Expert System to Diagnose Diseases in Corn Plant Using Mobile-Based Forward Chaining and Certainty Factor Methods. Tech Rom J Appl Sci Technol. **16**, 193–200 (2023). https://doi.org/10.47577/technium.v16i.9980
17. K. I. Payne, Warbot: The Dawn of Artificially Intelligent Conflict. (2021).
18. K. Thakur, A. S. K. Pathan, S. Ismat, Artificial Intelligence Technology. 45–77 (2023). https://doi.org/10.1007/978-3-031-27765-8_2
19. Y. Wang, NLP-driven integration of electrophysiology and traditional Chinese medicine for enhanced diagnostics and management of postpartum pain. SLAS Technol. **32**, 100267 (2025). https://doi.org/10.1016/j.slast.2025.100267
20. C. M. Liapis, A. Karanikola, S. Kotsiantis, Enhancing Sentiment Analysis with Distributional Emotion Embeddings. Neurocomputing. **634**, 129822 (2025). https://doi.org/10.1016/j.neucom.2025.129822
21. T. K. N. Narcisse, M. Sadouanouan, K. Malanno, B. K. K. Norbert, O. O. Germain, An Intelligent Research Environment on Cotton Diseases and Pests Based on a Cotton Phytosanitary Surveillance Ontology OntoSYSPARCOTCI. Procedia Comput Sci. **237**, 866–873 (2024). https://doi.org/10.1016/j.procs.2024.05.175
22. S. A. Aula, T. A. Rashid, FOX-TSA Hybrid Algorithm: Advancing for Superior Predictive Accuracy in Tourism-Driven Multi-Layer Perceptron Models. Syst Soft Comput. **6**, 200178 (2024). https://doi.org/10.1016/j.sasc.2024.200178
23. H. Abdulrahim, S. M. Alshibani, O. Ibrahim, A. A. Elhag, Prediction OPEC Oil Price Utilizing Long Short-Term Memory and Multi-Layer Perceptron Models. Alexandria Eng J. **110**, 607–612 (2025). https://doi.org/10.1016/j.aej.2024.10.057
24. Reinforcement Learning in Healthcare: Optimizing Treatment Strategies Dynamic Resource Allocation and Adaptive Clinical Decision-Making. Int J Comput Appl Technol Res. (2025). https://doi.org/10.7753/IJCATR1103.1007
25. Z. Morić, V. Dakić, S. Urošev, An AI-Based Decision Support System Utilizing Bayesian Networks for Judicial Decision-Making. Systems. **13**, 2, 131 (2025). https://doi.org/10.3390/systems13020131
26. B. P. Bhuyan, A. Ramdane-Cherif, T. P. Singh, R. Tomar, Rule-Based Systems and Expert Systems. 63–85 (2025). https://doi.org/10.1007/978-981-97-8171-3_5
27. S. F. Ahmed, M. S. Bin Alam, M. Hassan et al., Deep Learning Modelling Techniques: Current Progress Applications Advantages and Challenges. Artif Intell Rev. **56**, 11, 13521–13617 (2023). https://doi.org/10.1007/s10462-023-10466-8
28. H. GN, R. Siautama, I. A. AC, D. Suhartono, Extractive Hotel Review Summarization Based on TF-IDF and Adjective-Noun Pairing by Considering Annual Sentiment Trends. Procedia Comput Sci. **179**, 558–565 (2021). https://doi.org/10.1016/j.procs.2021.01.040
29. J. Zhou, Z. Ye, S. Zhang, Z. Geng, N. Han, T. Yang, Investigating Response Behavior through TF-IDF and Word2Vec Text Analysis: A Case Study of PISA 2012 Problem-Solving Process Data. Heliyon. **10**, 16, e35945 (2024). https://doi.org/10.1016/j.heliyon.2024.e35945
30. K. Thapa, M. Kinali, S. Pei, A. Luna, Ö. Babur, Strategies to Include Prior Knowledge in Omics Analysis with Deep Neural Networks. Patterns. **6**, 3, 101203 (2025). https://doi.org/10.1016/j.patter.2025.101203
31. K. Sandhu, D. N. Lozada, Z. Zhang, M. Pumphrey, A. Carter, Deep Learning for Predicting Complex Traits in Spring Wheat Breeding Program. Front Plant Sci. **11**, 613325 (2021). https://doi.org/10.3389/fpls.2020.613325
32. O. A. Montesinos López, A. Montesinos López, J. Crossa, Fundamentals of Artificial Neural Networks and Deep Learning. 379–425 (2022). https://doi.org/10.1007/978-3-030-89010-0_10
33. A. S. Shirazi, I. Frigaard, SlurryNet: Predicting Critical Velocities and Frictional Pressure Drops in Oilfield Suspension Flows. Energies. **14**, 1263 (2021). https://doi.org/10.3390/en14051263
34. Y. Luo, C. Lu, TF-IDF Combined Rank Factor Naive Bayesian Algorithm for Intelligent Language Classification Recommendation Systems. Syst Soft Comput. **6**, 200136 (2024). https://doi.org/10.1016/j.sasc.2024.200136
35. M. Liang, T. Niu, Research on Text Classification Techniques Based on Improved TF-IDF Algorithm and LSTM Inputs. Procedia Comput Sci. **208**, 460–470 (2022). https://doi.org/10.1016/j.procs.2022.10.064
36. U. Sirisha, P. N. Srinivasu, P. Padmavathi et al., Dynamic Multi-Layer Perceptron for Fetal Health Classification

Using Cardiotocography Data. Comput Mater Contin. **80**, 2, 2301–2330 (2024). https://doi.org/10.32604/cmc.2024.053132

37. F. Han, X. Fan, P. Long et al., Optimized Graph Neural Network–Multilayer Perceptron Fusion Classifier for Metastatic Prostate Cancer Detection in Western and Asian Populations. Asian J Urol. (2025). https://doi.org/10.1016/j.ajur.2024.09.009
38. E. Çakan, V. Rodoplu, C. Güzeliş, Data Fusion Integrated Network Forecasting Scheme Classifier via Multi-Layer Perceptron Decomposition Architecture. Internet of Things. **28**, 101341 (2024). https://doi.org/10.1016/j.iot.2024.101341
39. S. Narteni, V. Orani, E. Ferrari, D. Verda, E. Cambiaso, M. Mongelli, Explainable Evaluation of Generative Adversarial Networks for Wearables Data Augmentation. Eng Appl Artif Intell. **145**, 110133 (2025). https://doi.org/10.1016/j.engappai.2025.110133
40. F. J. Moreno-Barea, J. M. Jerez, L. Franco, Improving Classification Accuracy Using Data Augmentation on Small Data Sets. Expert Syst Appl. **161**, 113696 (2020). https://doi.org/10.1016/j.eswa.2020.113696
41. S. De, D. K. Sanyal, I. Mukherjee, Fine-Tuned Encoder Models with Data Augmentation Beat ChatGPT in Agricultural Named Entity Recognition and Relation Extraction. Expert Syst Appl. 127126 (2025). https://doi.org/10.1016/j.eswa.2025.127126
42. A. Das, F. Pathan, J. R. Jim, M. M. Kabir, M. F. Mridha, Deep Learning-Based Classification Detection and Segmentation of Tomato Leaf Diseases: A State-of-the-Art Review. Artif Intell Agric. **15**, 2, 192–220 (2025). https://doi.org/10.1016/j.aiia.2025.02.006
43. B. Singh, A. Gorbenko, A. Palczewska, H. Tawfik, Application of Machine Learning Techniques to Predict Teenage Obesity Using Earlier Childhood Measurements from Millennium Cohort Study. 55–60 (2023). https://doi.org/10.1145/3616131.3616139
44. A. Sumayli, Development of Advanced Machine Learning Models for Optimization of Methyl Ester Biofuel Production from Papaya Oil. Arab J Chem. 16, 7, 104833 (2023). https://doi.org/10.1016/j.arabjc.2023.104833
45. A. Tharwat, Classification Assessment Methods. Appl Comput Informatics. **17**, 1, 168–192 (2021). https://doi.org/10.1016/j.aci.2018.08.003
46. M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, R. Budiarto, Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking. IEEE Access. **8**, 90847–90861 (2020). https://doi.org/10.1109/ACCESS.2020.2994222
47. V. Malathi, M. P. Gopinath, M. Kumar, S. Bhushan, S. Jayaprakash, Enhancing the Paddy Disease Classification by Using Cross-Validation Strategy for Artificial Neural Network over Baseline Classifiers. J Sensors. (2023). https://doi.org/10.1155/2023/1576960
48. G. Canbek, T. Taskaya Temizel, S. Sagiroglu, PToPI: A Comprehensive Review Analysis and Knowledge Representation of Binary Classification Performance Measures. SN Comput Sci. **4**, 1, 13 (2022). https://doi.org/10.1007/s42979-022-01409-1
49. L. Li, T. T. Goh, D. Jin, How Textual Quality of Online Reviews Affect Classification Performance: A Case of Deep Learning Sentiment Analysis. Neural Comput Appl. **32**, 9, 4387–4415 (2020). https://doi.org/10.1007/s00521-018-3865-7
50. J. H. Cabot, E. G. Ross, Evaluating Prediction Model Performance. Surgery. **174**, 3, 723–726 (2023). https://doi.org/10.1016/j.surg.2023.05.023
51. J. Miao, W. Zhu, Precision–Recall Curve Classification Trees. Evol Intell. **15**, 3, 1545–1569 (2022). https://doi.org/10.1007/s12065-021-00565-2
52. R. Diallo, C. Edalo, O. O. Awe, Machine Learning Evaluation of Imbalanced Health Data: A Comparative Analysis of Balanced Accuracy MCC and F1 Score. 283–312 (2025). https://doi.org/10.1007/978-3-031-72215-8_12
53. S. Liu, J. Bao, Y. Lu, J. Li, S. Lu, X. Sun, Digital Twin Modeling Method Based on Biomimicry for Machining Aerospace Components. J Manuf Syst. **58**, 180–195 (2021). https://doi.org/10.1016/j.jmsy.2020.04.014