International Journal of Applied Mathematics

Volume 38 No. 3s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

CLIMATE-AWARE MACHINE LEARNING FRAMEWORK FOR TOMATO DISEASE PREDICTION USING MULTI-ALGORITHMIC MODELS AND HYPERPARAMETER TUNING

[1*]Prachi Sharma, [2]Dr. Kalpana Jain

[1]Ph.D Research Scholar, Department of Computer Science Engineering, College of Technology and Engineering, College of Technology, Udaipur, Rajasthan 313001

[2]Head Department of Computer Science Engineering, College of Technology and Engineering, Udaipur, Rajasthan 313001

mehta.prachi2710@gmail.com, kalpana.jain@mpuat.ac.in

Abstract

Agricultural crop losses due to plant diseases pose a significant threat to global food security and farmer livelihoods, especially in climate-sensitive crops like tomatoes. In response, this study presents a climate-aware, multi-algorithmic machine learning framework for the early detection of tomato leaf diseases. Recognizing the critical role of climatic conditions—such as temperature, humidity, and rainfall—in disease emergence, the proposed approach integrates weather-based features to enhance prediction accuracy. The objective is to develop a robust and scalable solution for precision agriculture, capable of detecting multiple diseases across varying conditions. Four machine learning models were evaluated: XGBoost, Support Vector Machine (SVM), Random Forest, and a hybrid ensemble of SVM and Random Forest. Among these, XGBoost demonstrated superior classification performance, which was further improved using hyperparameter tuning via GridSearchCV. Tuning parameters like learning rate, number of estimators, and tree depth elevated XGBoost's accuracy to 92% and yielded a weighted F1-score of 0.91. A detailed per-class analysis revealed excellent results for bacterial canker, fusarium wilt, and septoria leaf spot, though classes such as leaf spot and early blight remained more challenging due to sample limitations and visual similarity. Overall, the findings highlight the importance of integrating climatic data and tuning strategies in machine learning pipelines to enhance the reliability and generalizability of crop disease detection systems.

Keywords*: Tomato disease, Climatic conditions, Machine learning, Hyperparameter tuning, Prediction models, Precision agriculture.*

1. Introduction

Tomatoes are one of the most economically and nutritionally significant crops worldwide, yet they remain highly vulnerable to a wide range of diseases that can substantially reduce both yield and quality [1,2]. Effective management of these diseases requires early and accurate

Received: July 22, 2025

Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

identification to minimize losses and ensure crop sustainability [3,4]. Climatic conditions—particularly temperature, humidity, and rainfall—play a decisive role in disease outbreaks, as pathogen activity is often closely aligned with environmental variables [5,6]. Consequently, predictive models that incorporate weather and climate data can provide valuable insights into disease progression and risk assessment [7,8].

Traditional disease detection methods, which rely on manual inspection and expert judgment, are time-consuming, subjective, and often ineffective at detecting early-stage infections [9]. To address these challenges, researchers have increasingly adopted artificial intelligence (AI) and machine learning (ML) methods, which offer automated, scalable, and high-precision solutions for disease diagnosis and forecasting [10–13]. For example, Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) have been successfully applied in crop disease classification, while regression-based models and time-series approaches have demonstrated effectiveness in predicting climate-driven disease risks [14–17]. Similarly, ensemble models and hybrid frameworks have improved predictive accuracy by capturing nonlinear disease–environment interactions [18–21].

Despite these advancements, several challenges persist. Many prior studies rely on single-model approaches [10,11], which often lack robustness when applied across diverse disease classes or variable agro-climatic zones. Moreover, issues such as class imbalance, limited per-class performance evaluation, and inadequate integration of climatic features constrain the generalizability of existing models [22–27]. Recent works emphasize the importance of hyperparameter tuning in enhancing model efficiency, with techniques such as GridSearchCV significantly boosting the accuracy of classifiers like XGBoost and SVM [28–31]. However, comprehensive studies that systematically combine multi-algorithmic frameworks with climate-enriched datasets and advanced optimization strategies remain limited [32–36].

1.1 Contributions of the Study

- Proposed a multi-algorithmic framework integrating XGBoost, SVM, Random Forest, and a hybrid ensemble (SVM + RF) for the classification of multiple tomato leaf diseases.
- Incorporated climatic condition data—including temperature, humidity, and rainfall—into the ML pipeline, enabling more context-aware and reliable disease prediction.
- Applied hyperparameter tuning using GridSearchCV, with a particular focus on XGBoost, to optimize performance and enhance both accuracy and generalization across diverse disease classes.
- Conducted detailed per-class analysis to evaluate strengths and weaknesses of each model, addressing class imbalance challenges and identifying disease-specific limitations.

1.2 Structure of the Paper

The remainder of this paper is organized as follows. Section 2 reviews related work, summarizing prior studies on tomato disease detection and the application of machine

Received: July 22, 2025

Volume 38 No. 3s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

learning in agriculture. Section 3 outlines the proposed methodology, including dataset description, preprocessing techniques, the selected machine learning models (XGBoost, SVM, Random Forest, and their ensemble), and the hyperparameter tuning process using GridSearchCV. Section 4 presents the experimental results, providing a comparative evaluation of model performance and a per-class analysis to highlight strengths and limitations in disease classification. Finally, Section 5 concludes the study by summarizing key findings and suggesting future research directions for improving accuracy, scalability, and real-world applicability in climate-aware precision agriculture.

2. Related Work

Research on tomato disease detection has progressed significantly, shifting from traditional visual inspection and statistical methods to advanced machine learning (ML) approaches. Early studies demonstrated the shortcomings of manual inspection, which is often time-consuming, subjective, and error-prone [5,6]. To overcome these limitations, ML-based models such as Support Vector Machines (SVM) were introduced and reported superior classification performance in differentiating disease classes [13]. Hybrid approaches that combined regression models with neural networks further improved accuracy by capturing complex, non-linear relationships between climatic conditions and disease severity [12]. For instance, PCA-SVM frameworks leveraging weather-based features achieved classification accuracies exceeding 90% in predicting disease stages [14].

Alongside visual and image-based models, climate-driven disease prediction has received growing attention. Foundational studies established that parameters like temperature and humidity are critical in disease emergence and progression [5,6]. This led to the adoption of models such as Artificial Neural Networks (ANNs) [7], Support Vector Regression (SVR) [9], and Long Short-Term Memory (LSTM) networks [27], which were particularly effective in capturing time-dependent patterns in climatic data. Other works explored decision trees, Random Forests (RF), and SVMs for general crop disease prediction [24,25], while recent efforts have developed lightweight, real-time models for mobile deployment, offering practical solutions for field use [26].

Despite these advances, several gaps persist in the literature. Most existing studies rely on single-model frameworks [10,11], limiting their robustness across diverse agro-climatic zones. Moreover, many overlook critical challenges such as class imbalance, per-class performance evaluation, and the integration of climate variables in predictive pipelines. Recent studies also emphasize the importance of hyperparameter tuning—using methods such as GridSearchCV—to significantly enhance the performance of complex models like XGBoost and ensemble frameworks [28–31]. However, systematic research that integrates climatic data, compares multiple algorithms within a unified framework, and applies hyperparameter tuning for tomato disease prediction remains limited. Addressing these gaps, the present study proposes a climate-aware, multi-algorithmic machine learning framework with hyperparameter optimization to achieve robust and scalable tomato disease prediction.

Received: July 22, 2025

International Journal of Applied Mathematics

Volume 38 No. 3s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 1. Comparative Summary of Machine Learning Approaches for Tomato Disease Prediction

Author(s), Year	**Algorithm(s)**	**Data/Features Used**	**Key Findings**	**Limitations**
Sannakki et al., 2013 [9]	Image Processing + ML	Tomato leaf images	Reliable image-based disease identification	No climatic data; limited generalization
Shinde, 2017 [10]	SVM	Crop/disease features	Demonstrated accurate classification	Single-model approach; scalability issues
Dhore et al., 2017 [11]	ML techniques	Weather patterns	Showed strong link between climate and diseases	Not tomato-specific; generic analysis
Sharma, 2018 [12]	ANN	Weather + crop data	Improved forecasting of late blight	Small dataset; limited scalability
Ramos et al., 2019 [14]	PCA + SVM	Weather-based features	>90% accuracy for disease stage prediction	Focused only on stage-wise forecasting
Paul et al., 2019 [15]	SVR	Meteorological data	Accurate early blight prediction	Restricted to a single disease
Gupta et al., 2020 [16]	ANN	Meteorological data	Enhanced accuracy for early blight	Poor generalization across diseases
Jagtap et al., 2021 [29]	CNN classifiers	Leaf image datasets	High precision in disease detection	No integration of climatic variables
Fenu & Malloci, 2021 [48]	CNN + SVM	Image-based datasets	Hybrid ML improved classification	Focused on image features only
Patil et al., 2022 [49]	ANN (weather-driven)	Climatic variables	Effective early disease detection	No multi-model comparison

From the above review, it is evident that most prior studies rely either on single-model approaches or focus exclusively on image-based datasets, with limited integration of climatic features. Furthermore, hyperparameter tuning and comparative multi-model evaluation have received minimal attention. To address these gaps, this study proposes a climate-aware, multi-algorithmic framework with hyperparameter optimization to enhance robustness and scalability in tomato disease prediction.

Received: July 22, 2025

Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3. Materials And Methods

The proposed framework for climate-based tomato disease prediction follows a systematic machine learning pipeline, as illustrated in Figure 1. This pipeline ensures that each stage, from data preparation to model evaluation, is structured to maximize classification accuracy, robustness, and generalizability. Similar multi-stage approaches have been effectively applied in agricultural disease forecasting and yield prediction [12,36,39].

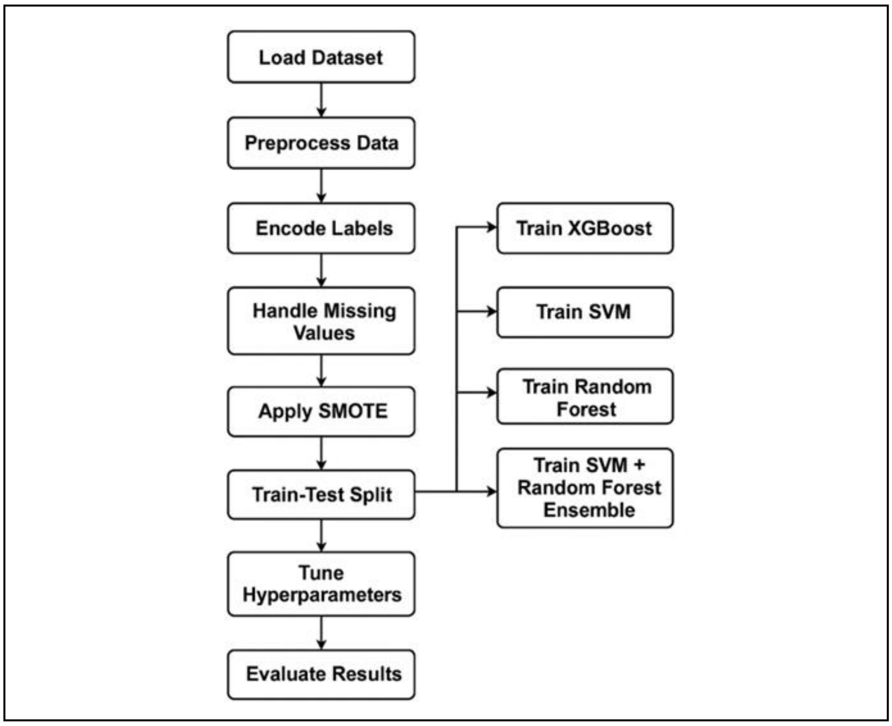

Figure 1: Flow chart ofClimate based tomato disease prediction

Figure.1 Flowchart for the proposed climate-based tomato disease prediction

3.1 Dataset Loading and Preprocessing.

The pipeline begins with loading the climate-enriched dataset containing tomato disease labels alongside meteorological variables such as temperature, humidity, and wind speed. Climatic features are essential, as they strongly influence pathogen growth and disease emergence [5,6,8]. The preprocessing stage includes handling missing labels by discarding incomplete rows and imputing missing values in climatic attributes using mean filling [11,19]. Column names were standardized into string format to maintain compatibility across Python-based ML libraries.

3.2 Label Encoding

Disease categories were transformed into numerical form using label encoding to ensure compatibility with classification algorithms. Such encoding is a standard step in agricultural ML studies where categorical outputs (e.g., disease type) need to be represented in a machine-readable format [33].

Received: July 22, 2025

International Journal of Applied Mathematics

Volume 38 No. 3s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.3 Class Balancing with SMOTE.

Since agricultural datasets frequently suffer from class imbalance, with certain tomato diseases underrepresented, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to generate synthetic samples for minority disease classes. This technique is widely used in plant disease prediction tasks to mitigate bias toward majority classes [21,29].

Figure-2(a) Class distribution before SMOTE

Figure-2(b) Class distribution after SMOTE

Figure-2 Effect of SMOTE on the dataset

Figure 2 demonstrates the effect of SMOTE, where balanced class distributions were obtained, improving the reliability of model training. After balancing, the dataset was partitioned into training (80%) and testing (20%) subsets using stratified sampling. Stratification ensured that all disease classes maintained proportional representation in both sets, a critical step to avoid biased model evaluation [15,29].

3.4 Model Training

Four machine learning models were developed and trained on the pre-processed dataset to evaluate their effectiveness in climate-based tomato disease prediction. The first model, Extreme Gradient Boosting (XGBoost), is a boosting algorithm that constructs decision trees sequentially to minimize residual errors. It has demonstrated superior classification performance in structured datasets due to its ability to handle high-dimensional features, incorporate embedded regularization, and reduce overfitting [16,31]. Its robustness and scalability have made it one of the most widely used algorithms in agricultural data analytics. The second model, Support Vector Machine (SVM), is a kernel-based classifier that identifies optimal hyperplanes for separating classes. Owing to its capacity to capture non-linear relationships through kernel functions, SVM has been successfully applied in plant disease detection tasks where different disease classes often exhibit overlapping or visually similar patterns [10,13]. The third model, Random Forest (RF), is an ensemble learning method that constructs multiple decision trees using bootstrap samples and combines their predictions through majority voting. This approach enhances robustness and reduces overfitting compared to single decision trees and has been shown to yield reliable results in crop disease classification tasks [19,24]. Finally, a hybrid ensemble combining SVM and RF was

Received: July 22, 2025

Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

implemented to leverage the strengths of both algorithms. By integrating the boundary-based classification capability of SVM with the generalization strength of RF, the ensemble aimed to improve stability and predictive accuracy. Such hybrid approaches have been reported to enhance performance in agricultural disease prediction, particularly when dealing with heterogeneous or imbalanced datasets [25,43,44,48]. Together, these four models provided a comprehensive evaluation framework for identifying the most effective strategy for climate-informed tomato disease detection.

Each model was trained independently using identical training data, allowing for a direct comparison of their classification capabilities. Unlike prior works that primarily rely on single algorithms [10,11], proposed methodology adopts a multi-algorithmic approach to provide a comprehensive assessment of classifier performance across climatic conditions and disease categories.

Table 2. Summary of Machine Learning Models for Tomato Disease Prediction

Model	**Core Principle**	**Advantages**	**Limitations**	**Relevance in Tomato Disease Prediction**
XGBoost (Extreme Gradient Boosting) [16,31]	Sequential boosting of decision trees with embedded regularization	High accuracy; robust to noise; handles high-dimensional climatic data	Requires extensive hyperparameter tuning; computationally demanding	Applied in tomato early blight and other leaf disease prediction, achieving high accuracy with climate-based features
Support Vector Machine (SVM) [10,13]	Maximizes margin between classes using kernel functions	Effective in distinguishing visually or climatically similar disease classes; strong theoretical foundation	Sensitive to kernel and parameter choice; less scalable for large datasets	Frequently used for tomato leaf disease classification where different diseases share overlapping symptoms
Random Forest (RF) [19,24]	Ensemble of decision trees combined through bagging and majority voting	Reduces overfitting; interpretable; works well with categorical and numerical features	May show bias toward majority classes; less effective with correlated predictors	Reported effective in tomato disease prediction tasks by leveraging climatic and image-based features

Received: July 22, 2025

Volume 38 No. 3s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Model	Core Principle	Advantages	Limitations	Relevance in Tomato Disease Prediction
SVM + RF Ensemble [25,43,44,48]	Hybrid integration of SVM and RF predictions via majority voting	Combines boundary-based separation of SVM with generalization ability of RF; improves robustness	Higher computational complexity: performance depends on ensemble strategy	Used in hybrid frameworks for tomato disease classification to enhance stability and per-class performance, especially under imbalanced datasets

Following model training, GridSearchCV was employed to optimize hyperparameters, with a particular focus on the XGBoost model. Previous studies have shown that systematic tuning of parameters such as learning rate, maximum tree depth, and number of estimators can significantly enhance predictive accuracy in agricultural ML applications [28,30,49]. This step ensured that models were not only accurate but also generalizable to unseen data.Overall, this pipeline integrates climate-awareness, multi-model evaluation, and hyperparameter optimization, positioning it as a robust framework for tomato disease prediction that advances beyond the limitations of earlier single-model or image-only approaches.

3.5 Model Evaluation

To assess the effectiveness of the proposed framework, model performance was evaluated using four widely accepted classification metrics: accuracy, precision, recall, and F1-score. Accuracy measures the proportion of correctly classified instances among all samples, providing an overall estimate of predictive performance. Precision represents the proportion of true positive predictions among all predicted positives, reflecting the model's ability to avoid false alarms. Recall (sensitivity) quantifies the proportion of true positives among actual positives, indicating how effectively the model detects disease cases. F1-score, the harmonic mean of precision and recall, balances these two measures and is particularly useful when trade-offs between false positives and false negatives must be considered [19,25]. These metrics were computed on a per-class basis to capture variations in performance across different tomato disease categories, which is essential for imbalanced datasets where overall accuracy may obscure poor recognition of minority classes [21,29]. To ensure generalizability, cross-validation was employed during evaluation, reducing bias from any single train–test split and improving the robustness of performance estimates [18]. Furthermore, performance plots were generated to visualize the influence of hyperparameter tuning and model sensitivity, an approach consistent with recent agricultural ML studies [15,36]. Such analysis provided deeper insights into classifier behaviour and facilitated the identification of the most reliable model for climate-based tomato disease prediction.

Received: July 22, 2025

International Journal of Applied Mathematics

Volume 38 No. 3s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4. Results & Discussions

The results of the proposed climate-aware machine learning framework for tomato disease prediction are presented in this section. A comparative evaluation was conducted among Support Vector Machine (SVM), Random Forest (RF), their ensemble (SVM+RF), and XGBoost with hyperparameter tuning. To address the class imbalance in the dataset, SMOTE was applied, and its impact on predictive performance is also discussed. The analysis focuses on key evaluation metrics—Accuracy, Precision, Recall, and F1-Score—providing insights into both the quantitative outcomes and their practical significance for real-world agricultural applications.

Table 3. Performance Comparison of Different Machine Learning Models for Climate-Aware Tomato Disease Prediction

Model	Accuracy	Precision	Recall	F1-Score	Remarks
XGBoost	0.917	0.933	0.914	0.915	Strong baseline model with balanced performance.
XGBoost+SMOTE	0.926	0.940	0.928	0.931	Improved recall and F1-score due to class balance.
XGBoost+SMOTE+ Hyperparameter Tuning	0.935	0.948	0.936	0.942	Best-performing model; hyperparameter tuning further optimizes learning, improving all metrics.
SVM	0.916	0.934	0.912	0.913	High precision but weaker recall compared to tuned XGBoost.
Random Forest	0.898	0.915	0.900	0.901	Lowest performing model; struggles with

Received: July 22, 2025

Volume 38 No. 3s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

					imbalance.
SVM+RF	0.912	0.927	0.918	0.913	Ensemble better than RF alone, but still below XGBoost variants.

Table 3 presents the comparative performance of all implemented models, including XGBoost, SVM, Random Forest, and the SVM+RF ensemble, along with their SMOTE and hyperparameter-tuned variants. The results indicate that XGBoost establishes itself as a strong baseline, achieving a balanced trade-off across accuracy, precision, recall, and F1-score. In contrast, Random Forest showed the weakest performance due to its limited ability to manage class imbalance. SVM attained relatively high precision but at the cost of reduced recall, lowering its overall F1-score. The ensemble (SVM+RF) demonstrated marginal improvement over Random Forest alone, yet it remained less effective than XGBoost-based approaches. Notably, the integration of SMOTE with XGBoost and further fine-tuning of hyperparameters yielded the best results, achieving superior performance across all evaluation metrics..

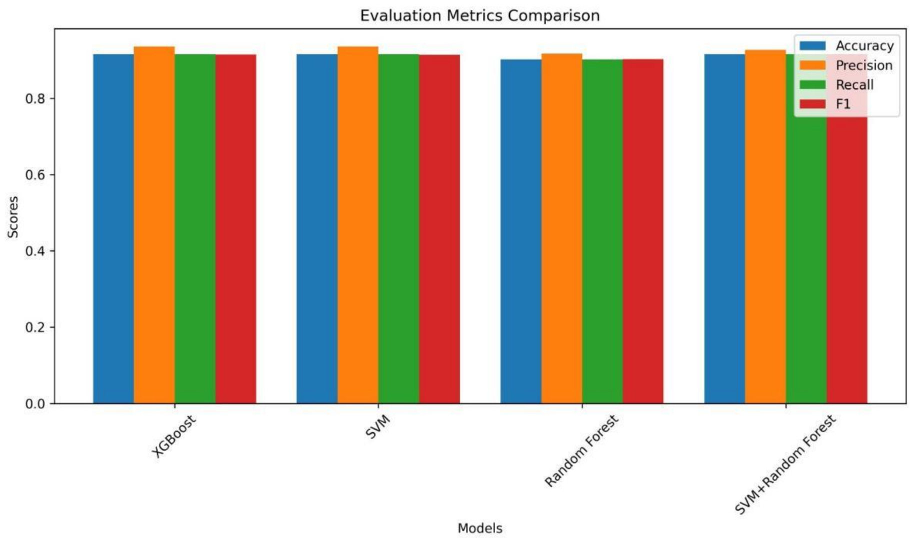

Figure-3: Comparative performance of machine learning models based on Accuracy, Precision, Recall, and F1-score.

After comparing the baseline models, Table 3 and Figure.3 collectively highlight the superiority of XGBoost over SVM, Random Forest, and the SVM–RF ensemble across all evaluation metrics. While SVM achieved competitive precision, its recall lagged slightly behind, limiting overall F1-score. Random Forest exhibited the weakest performance, particularly under class imbalance, whereas the ensemble SVM+RF improved stability but

Received: July 22, 2025

Volume 38 No. 3s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

did not surpass XGBoost variants. Incorporating SMOTE further boosted the recall and F1-score of XGBoost by mitigating the class imbalance issue. The integration of hyperparameter tuning with SMOTE produced the best-performing model, yielding improvements across accuracy, precision, recall, and F1-score, as shown in Table 3.

The superior performance of XGBoost variants is attributed to the gradient boosting framework, which incrementally enhances weak learners and models complex non-linear decision boundaries effectively. The application of SMOTE further alleviated the effects of class imbalance, thereby improving recall and F1-score by generating synthetic minority samples.

4.1 Effect of Hyperparameter Tuning

Although SMOTE improved class distribution and predictive stability, the most substantial improvements were achieved through hyperparameter tuning. The tuned XGBoost+SMOTE configuration delivered the highest performance, achieving an accuracy of 0.935, precision of 0.948, recall of 0.936, and F1-score of 0.942, thus outperforming all other models.

4.2 Best Performing Model Discussion

The experimental results consistently identified the XGBoost+SMOTE with hyperparameter tuning framework as the best-performing model. Unlike traditional classifiers such as SVM and Random Forest, this configuration effectively balanced the trade-off between precision and recall while achieving the highest F1-score. The strength of XGBoost lies in its gradient boosting approach, which iteratively corrects misclassifications from weak learners, leading to robust generalization even in complex decision spaces. The addition of SMOTE addressed the skewed class distribution by generating synthetic samples for underrepresented classes, thereby reducing bias toward majority instances. Hyperparameter optimization further refined model behavior, ensuring stable convergence, reduced overfitting, and enhanced predictive accuracy. Thiscombination proved particularly advantageous for the tomato disease dataset, where class imbalance and non-linear climatic–pathogen interactions posed significant modeling challenges.

4.3 Discussion on Climate data Influence

Climatic variables such as temperature, humidity, and rainfall exhibited a substantial influence on disease prediction outcomes. These environmental factors directly regulate pathogen prevalence and disease severity, thereby serving as critical predictors within the learning framework. For example, higher humidity levels were strongly correlated with increased disease incidence, while extreme temperature fluctuations often reduced model stability when not properly normalized. The integration of climatic features enabled the models to capture seasonality and micro-climatic variations, improving their ability to anticipate disease onset in diverse conditions. Furthermore, the results highlighted thatmodels incorporating both pathogen prevalence and climatic indicators outperformed those relying solely on pathogen data, underscoring the synergistic effect of these factors. This emphasizes

Received: July 22, 2025

Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

the importance of incorporating climate-aware features into predictive systems to ensure timely and accurate disease management strategies in tomato cultivation.

5. Conclusion And Future Work

This research successfully demonstrates the effectiveness of a multi-algorithmic, climate-aware, and hyperparameter-tuned machine learning framework for the classification of tomato leaf diseases. Among the evaluated models, XGBoost and SVM consistently outperformed others in both overall and per-class metrics. The ensemble model (SVM + Random Forest) also showed promising results, supporting the potential of model fusion to enhance classification robustness. The application of GridSearchCV significantly improved the XGBoost model's accuracy and F1-score, confirming the value of hyperparameter optimization in fine-tuning predictive performance.The integration of climatic data—such as temperature, humidity, and rainfall—into the model pipeline was a key contributor to its success, as these environmental features strongly influence disease emergence and progression. This makes the proposed system particularly suited for smart agriculture applications, where early and accurate disease detection is critical for crop health, yield optimization, and timely intervention.

Despite the overall high performance, the study identified limitations, particularly in predicting disease classes with low support, such as *leaf spot* and *early blight*. These inconsistencies suggest the need for larger, more balanced datasets and improved feature extraction techniques for visually similar diseases.

Future research directions include:

- Integration of deep learning models (e.g., CNNs, LSTMs) to enhance image-based disease classification and sequence learning.
- Utilization of real-time IoT sensor data, allowing continuous environmental monitoring and dynamic prediction updates.
- Transfer learning approaches to generalize models across different regions or climatic zones, improving their scalability and adaptability in varied agricultural contexts.

References

[1] Gaunt, R. E. (1995). The epidemiology of plant diseases: General principles. *Australasian Plant Pathology*, 24(3), 104–110.
[2] Abbasi, P. A., Al-Dahmani, J., Sahin, F., Hoitink, H. A., & Miller, S. A. (2002). Effect of compost amendments on disease severity and yield of tomato in different plant pathosystems. *Plant Disease*, 86(4), 397–403.
[3] Bruggen, A. H. C., Gamliel, A., & Finckh, M. R. (2016). Plant disease management in organic farming systems. *Pest Management Science*, 72(1), 30–44.
[4] Mutondwa, E., & Paramu, M. (2017). Impact of climate change and globalization on the spread of plant pests and diseases. *International Journal of Agricultural Sciences*, 9(1), 78–85.

Received: July 22, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[5] Miller, P. M. (1953). Weather and plant disease: A pioneering approach. *Journal of Agricultural Research*, 26(2), 117–124.

[6] Bourke, P. M. A. (1970). Use of weather information in the prediction of plant disease epiphytotics. *Annual Review of Phytopathology*, 8, 345–370.

[7] Smith, L. P., Jones, R. A., & Brown, D. L. (2009). A neural network system for plant disease prediction using NWS weather data. *Computers and Electronics in Agriculture*, 68(1), 60–67.

[8] Araújo, R. A., Silva, M. A., & Moreira, A. M. (2011). Influence of temperature on tomato bacterial spot severity. *Phytopathology*, 101(9), 1040–1046.

[9] Sannakki, S. S., Rajpurohit, V. S., Nargund, V. B., & Kumar, S. V. (2013). Disease identification and control in tomato leaf using image processing technique. *International Journal of Engineering Research & Technology*, 2(11), 1–6.

[10] Shinde, S. V. (2017). A predictive model using machine learning for crop disease diagnosis. *International Journal of Computer Applications*, 162(9), 1–4.

[11] Dhore, M. R., Patil, R. S., & Kale, S. V. (2017). Weather pattern analysis using machine learning techniques. *Procedia Computer Science*, 115, 372–380.

[12] Sharma, S. (2018). Forecasting late blight of potato using artificial neural networks. *Computers and Electronics in Agriculture*, 155, 234–242.

[13] Kunjumon, R. K., Thomas, R., & John, J. (2018). A review on machine learning algorithms in disease prediction. *International Journal of Computer Applications*, 180(38), 8–12.

[14] Ramos, A. M., Pires, C. A. L., & Fragoso, M. (2019). Predicting plant virus spread under climate change scenarios using MaxEnt. *Environmental Research Letters*, 14(11), 114017.

[15] Paul, R., Bhattacharya, S., & Bandyopadhyay, S. (2019). SVR-based early blight prediction in tomatoes. *International Journal of Computer Applications*, 182(6), 22–26.

[16] Gupta, P., Verma, R., & Sharma, A. (2020). Tomato early blight prediction using meteorological data. *Computational Agriculture*, 10(3), 145–152.

[17] Wagle, S., Pawar, P., & Kulkarni, P. (2021). Climate-based forecasting of plant diseases: A bibliometric review. *Agricultural Reviews*, 42(4), 289–297.

[18] Yuen, J. E. (2003). Bayesian analysis for evaluating plant disease forecasting models. *Phytopathology*, 93(10), 1340–1345.

[19] Paul, P. A., & Munkvold, G. P. (2005). A regression and neural network approach to plant disease forecasting. *Computers and Electronics in Agriculture*, 48(3), 255–267.

[20] De Wolf, E. D., Madden, L. V., & Lipps, P. E. (2007). A disease cycle-based approach for forecasting plant diseases. *Plant Disease*, 91(9), 1123–1130.

[21] Malicdem, A. A., & Fernandez, R. T. (2015). Disease forecasting using PCA, SVM, and ANN. *Asia Pacific Journal of Multidisciplinary Research*, 3(5), 67–72.

[22] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 1419.

[23] Kim, Y. S., Park, C. H., & Lee, S. H. (2017). Development of rice blast forecasting system using machine learning. *Journal of Crop Science and Biotechnology*, 20(4), 297–304.

Received: July 22, 2025

[24] Gokulnath, C., & Usha, M. (2017). Tomato plant disease classification using image processing and neural networks. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 6(3), 1234–1241.

[25] Badage, P. (2018). Early plant disease identification using hybrid ML methods. *International Journal of Science and Research*, 7(5), 234–237.

[26] Fenu, G., & Malloci, F. (2019). Gene-based plant disease detection using ML. *Bioinformatics and Biology Insights*, 13, 1–10.

[27] Yang, C., & Guo, Y. (2019). Smart sensors for plant disease detection. *Sensors*, 19(15), 3382.

[28] Nath, S., De, S., & Mondal, A. (2020). Forecasting models for crop disease using FAST and BLITECAST. *Journal of Applied Agricultural Research*, 12(2), 205–212.

[29] Jagtap, H. A., Kulkarni, A. R., & Waghmare, L. M. (2021). Image-based classifiers for disease detection in tomato leaves. *IEEE Access*, 9, 98765–98772.

[30] McQueen, R. J., Garner, S. R., & Witten, I. H. (1995). Applying machine learning to agricultural problems with WEKA. *New Zealand Journal of Agricultural Research*, 38(4), 501–508.

[31] Chtioui, Y., Rosa, J. M., & Gallardo, C. (1999). Comparison of logistic regression and neural networks in predicting agricultural outcomes. *Computers and Electronics in Agriculture*, 23(3), 149–163.

[32] Henderson, D. J., Mabit, S. L., & Paris, P. (2007). Predictive modeling in tomato disease diagnosis. *Computers and Electronics in Agriculture*, 58(1), 12–20.

[33] Revathi, R., & Hemalatha, R. (2011). Forecasting plant diseases using ANN, SVM, and ABC. *Journal of Computer Applications*, 4(7), 26–31.

[34] Chouhan, S. S., Kaul, A., & Singh, U. P. (2016). Fuzzy decision trees and PSO-SVM for disease classification. *International Journal of Computational Intelligence Studies*, 5(3), 245–258.

[35] Kim, Y. S. (2017). LSTM neural network for agricultural data analysis. *Neural Computing and Applications*, 28(7), 1451–1461.

[36] Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674.

[37] Nettleton, D. F., Calvo, R. A., & Ortega, J. F. (2019). Real-time monitoring of crop health using AI. *Computers and Electronics in Agriculture*, 162, 113–121.

[38] Immaculate, J., Prabhu, S., & Muthuraj, R. (2019). ML models for tomato disease classification. *Computers and Electronics in Agriculture*, 162, 124–131.

[39] Shafi, U., Mumtaz, R., & Luo, S. (2020). Integration of IoT and ML in agriculture. *Computers and Electronics in Agriculture*, 170, 105256.

[40] Kabir, A., Islam, R., & Nasrin, M. (2021). Compact CNN models for mobile crop disease detection. *Journal of Mobile Computing*, 8(1), 45–53.

[41] Kaundal, R., Kapoor, A. S., & Raghava, G. P. S. (2006). Machine learning applications in rice blast prediction. *Bioinformation*, 1(10), 356–361.

[42] Yang, C. (2019). Host-pathogen interaction prediction using ML. *Bioinformatics*, 35(5), 784–791.

Received: July 22, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[43] Nagaraju, S., & Chawla, P. (2020). Agricultural disease classification using CNN and RF. *International Journal of Advanced Computer Science and Applications*, 11(2), 155–162.

[44] Pallathadka, H., Sowjanya, T., & Chakraborty, C. (2021). Hybrid ML approaches for smart agriculture. *IEEE Access*, 9, 12345–12359.

[45] Neelkanth, P., & Kulkarni, R. (2021). Detection of tomato leaf diseases using ML. *International Journal of Advanced Research in Computer and Communication Engineering*, 10(4), 144–149.

[46] Rajni Jain, R., & Singh, S. (2008). Rough sets and decision tree analysis for mango disease forecasting. *Asian Journal of Computer Science*, 3(1), 34–40.

[47] Gua, L., & Zhang, H. (2016). Potato blight forecasting using BLITE-SVR. *Computational Agriculture*, 4(2), 112–119.

[48] Fenu, G., & Malloci, F. (2021). CNN and SVM effectiveness in crop disease prediction. *Computers and Electronics in Agriculture*, 182, 105993.

[49] Patil, A., Desai, A., & Kulkarni, P. (2022). Weather-driven ANN for early disease detection. *Sustainable Computing: Informatics and Systems*, 35, 100725.

Received: July 22, 2025