Turkish Journal of Agriculture - Food Science and Technology, 14(7): 1963-1970, 2026
DOI: https://doi.org/10.24925/turjaf.v14i7.1963-1970.8826

Turkish Journal of Agriculture - Food Science and Technology

Available online, ISSN: 2148-127X | www.agrifoodscience.com | Turkish Science and Technology Publishing (TURSTEP)

The Classification of Germination Parameters Under Different Salinity Conditions Using Machine Learning Approaches

Mustafa Turan Arslan[1,a], Gülay Zulkadir[2,b,*]

[1]*Hatay Mustafa Kemal University, Kırıkhan Vocational School, Department of Computer Technology, 31440, Hatay, Türkiye*
[2]*Mersin University, Applied Technology and Management School of Silifke, Department of Organic Agriculture and Management, 33940, Mersin, Türkiye*
[*]*Corresponding author*

ARTICLE INFO

Research Article

Received : 19.02.2026
Accepted : 29.03.2026

Keywords:

Classification
Abiotic stress
Zea mays L.
Genotype selection
Machine learning algorithms

ABSTRACT

Soil salinity is a major abiotic stress affecting maize (Zea mays L.) production, leading to reduced germination and yield. Traditional breeding methods are resource-intensive and limited by environmental variability. This study explores machine learning (ML) algorithms to classify salinity stress levels based on germination parameters, aiming to enhance breeding efficiency for salt-tolerant genotypes. Germination data from 39 maize genotypes were collected under three salinity conditions (0 mM, 100 mM, 200 mM NaCl). Parameters included T10, T50, T90, T10-90, MGT, and Gmax. Data preprocessing involved normalization and feature selection using Correlation-based Feature Selection (CFS). Five ML models BayesNet, Decision Table, Bagging, J48, and Random Forest were evaluated using 10-fold cross-validation, with performance metrics of accuracy, precision, and F-measure. Results: Using all features, Random Forest achieved the highest accuracy (92.02%), followed by BayesNet (89.74%). After CFS (removing genotype and T10-90), BayesNet performed best (89.74%), while Random Forest slightly declined (87.46%). All models exceeded 87% accuracy, demonstrating robust classification of salinity levels. Conclusions: ML models effectively classify salinity stress in maize germination, even without certain parameters, supporting rapid genotype selection. This approach offers a scalable tool for improving agricultural resilience against salinization.

[a] *mtarslan@mku.edu.tr* *https://orcid.org/0000-0001-5498-6571* | [b] *gulayzulkadir@gmail.com* *https://orcid.org/0000-0003-3488-4011*

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Introduction

Soil salinity ranks among the primary environmental stressors impacting crop yields globally, especially in arid and semi-arid areas. Research suggests that salinity affects 20-50% of irrigated lands, with some studies indicating that 33% of such lands worldwide face salinity and sodicity issues (Isik et al., 2023; Mohanavelu et al., 2021; Pitman & Läuchli, 2002). This environmental issue adversely affects plant growth, germination, and yield due to the toxic buildup of sodium (Na^+) and chloride (Cl^-) ions and osmotic stress, posing a significant threat to food security and sustainable agriculture (Collado et al., 2016; Isik et al., 2023). For maize (Zea mays L.), a crucial cereal crop, salinity severely hampers cultivation and leads to notable global production losses. Despite its vast genetic diversity, the limited understanding of salt tolerance mechanisms complicates breeding efforts (Guo et al., 2023; Ibrar et al., 2024; Zulkadir, 2025).

Elevated salinity levels lead to delayed germination, decreased viability, and low emergence rates by disrupting osmotic balance, ion homeostasis, and enzymatic functions (Khan et al., 2006). Previous research has shown that salinity stress extends germination parameters like T50 and MGT while decreasing Gmax. Sensitive maize varieties are significantly affected at NaCl concentrations above 100 mM, and even tolerant varieties are inhibited at 200 mM and higher (Khalid et al., 2023; Zulkadir, 2025).

Traditional agricultural methods depend on labor-intensive and time-consuming field trials for selecting genotypes under salt stress. However, these approaches have limited applicability due to strong genotype-by-environment (G×E) interactions (Atanda et al., 2021; Luo et al., 2017). At this juncture, machine learning algorithms offer significant advantages in plant breeding by effectively integrating high-throughput phenotyping and

genotyping data, thereby enhancing prediction accuracy and breeding efficiency.

Machine learning techniques can identify relevant traits from extensive genomic, phenotypic, and environmental datasets and develop robust prediction models using cross-validation methods. Genomic selection supported by machine learning reduces the need for extensive phenotyping, shortens breeding cycles, and improves the selection of superior genotypes (Tong & Nikoloski, 2021). Among the commonly used algorithms, Random Forest (RF) generally achieves high accuracy in yield prediction. Additionally, combinations of XGBoost and RF using soft voting techniques have shown superior performance with over 98% accuracy in some studies (Altaf et al., 2026; Maheswari & Ramani, 2023; Premasudha et al., 2022).

However, machine learning applications in sustainable agriculture still face significant generalization challenges due to data quality issues and high field variability, which limit their practical implementation across different conditions (Botero-Valencia et al., 2025; Chintakunta et al., 2023).

This study aims to assess the performance of several ML-based classification models (BayesNet, Decision Table, Bagging, J48, and Random Forest) using germination data from 39 different maize genotypes under various salinity levels. The goal is to enable rapid and accurate selection of salt-tolerant genotypes, thereby increasing the efficiency of maize breeding programs through metrics such as Accuracy, Precision, and F-Measure. This approach is expected to contribute to strengthening agricultural resilience against the growing threat of soil salinization due to climate change.

Materials and Methods

In this study, experimental data obtained to evaluate the germination performance of corn genotypes under different environmental conditions (no salinity, low salinity, and salinity) were used. Different corn genotypes were tested in the experiments under different environmental conditions. The basic steps of the study are shown in Figure 1. Pre-processing steps were first applied to the dataset, and then classification was performed using different machine learning methods.

Seeds from thirty-nine maize genotypes were selected based on homogeneity and viability. Germination tests were conducted in a controlled growth chamber at 25°C and 70% relative humidity. Salinity treatments were performed using NaCl solutions: 0 mM (salt-free, control), 100 mM (low salinity), and 200 mM (high salinity), reflecting common stress thresholds in the literature (Carpici et al., 2009; Khalid et al., 2023). For each treatment-genotype combination, 50 seeds were placed on filter paper in Petri dishes moistened with the relevant solution, and three replicates were performed per treatment. Germination was defined as root emergence >2 mm and monitored daily for 14 days. Parameters were calculated using standard formulas: T10, T50, T90 as cumulative germination percentages; T10-90 as T90 - T10; MGT as $\Sigma(Dn)/\Sigma n$ (D = days, n = number of seeds germinating per day); Gmax as the final germination percentage (ISTA, 2015).

Data Description

A new dataset was collected to analyze the germination performance of maize genotypes under different environmental conditions in this study. The dataset contains seven different germination parameters: genotype, Time (days) until 10% germination occurs (T10), Time (days) until 50% germination occurs (T50), Time (days) until 90% germination occurs (T90), Time (days) between 10% and 90% occurs (T10-90), Mean Germination Time (days) (MGT), and Maximum Germination Percentage (Gmax). These parameters were considered important biological indicators representing the germination dynamics of each environment. The parameters were used as input variables in the classification models. The original dataset consists of 351 samples. 117 samples were in a non-saline environment, 117 in a low-saline environment, and the remaining 117 in a saline environment. The classes in the dataset are well-balanced. Thus, all three classes are equally represented in the dataset. The attributes in the dataset are described in detail in Table 1.

Table 1 presents attribute name, type of attribute, and basic statistical data such as mean, standard deviation, minimum, and maximum of the parameters in the data set.

Pre-Processing

The purpose of data preprocessing is to increase the accuracy rates of machine learning methods by performing some statistical analyses on raw data in order to produce high-performance machine learning models. This process attempts to improve data quality by employing methods such as normalization and feature selection. In this study, the germination parameters were normalized by scaling them between 0 and 1. Furthermore, feature selection was performed to remove unnecessary and irrelevant features from the data. Table 2 indicates the mean and standard deviation values for each attribute after this process.

In this study, the performance of a dataset containing germination parameters in environments with varying salinity was evaluated using machine learning methods. The models used in this study include Bayesian Net, Naive Bayes, and Decision Tree, as well as a number of ensemble and tree-based methods such as Bagging, J48 Tree, and Random Forest. A 10-fold cross-validation technique was used to reliably measure the overall performance of the models. This technique allowed for accurate comparisons of classifier performance on the dataset.

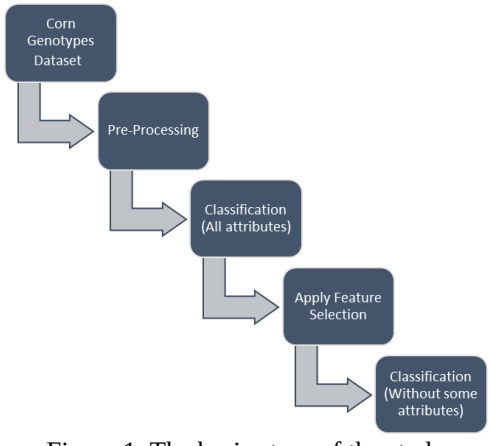

Figure 1. The basic steps of the study

Table 1. The basic statistical information of the dataset attributes.

Attribute Name	Type of Data	Min Value	Max Value	Mean Value	Standard deviation
Genotype	Categorical	1	39		-
T10	Numerical	0	11.3	5.308	2.582
T50	Numerical	0	12.5	7.096	2.845
T90	Numerical	0	13.7	9.496	3.009
T10-90	Numerical	0	8.6	4.188	1.37
MGT	Numerical	0	14	8.323	2.923
Gmax	Numerical	0	100	69.957	27.413

Table 2. Mean and Standard Deviation after Normalization Process.

Attribute	Mean	Standard Deviation
T10	0.47	0.229
T50	0.568	0.228
T90	0.693	0.22
T10-90	0.487	0.157
MGT	0.594	0.209
Gmax	0.7	0.274

Bayesian Network

The Bayesian Net method is based on Bayes' theorem. This method is a powerful and successful algorithm for dealing with high-dimensional data containing complex relationships (Condon & Revelle, 2014).

$$P(A/B) = \frac{P(B/A).P(A)}{P(B)} \quad (1)$$

P(A) denotes the probability of event A, P(B) represents the probability of event B occurring under all possible scenarios, and P(B|A) refers to the probability of observing event B given that event A has occurred.

Decision Table

A Decision Table is a classification algorithm that generates decisions by generating results corresponding to logical conditions in a table format. In this method, each row is assumed to represent a specific combination of conditions, and when these conditions are met, the relevant decisions are made. Conditions and outcomes are defined based on specific attributes, and the columns in the table represent possible scenarios (Gitmez et al., 2019; Keles & Keles, 2018). This model is widely used in the field of machine learning because it simplifies complex decision-making processes.

$$f(c_1, \dots, c_n) = \sum_{k=1}^{r} A_k . I_k(c_1, \dots, c_n) \quad (2)$$

Here, I_k denotes the indicator function that specifies whether rule k is satisfied, A_k represents the action(s) to be applied, and c_i corresponds to the value of the i-th condition for rule k.

Bagging

Bagging is a classification algorithm that combines multiple versions of the same model. This method uses majority voting for categorical predictions, while for numerical data, it averages these versions of the same model. These multiple versions are constructed using learning sets generated by random sampling from the training data (Breiman, 1996).

$$f_{bag}(x) = \frac{1}{B} \quad (3)$$

Here, h_b denotes the trained models and B denotes the number of models.

J48

Bagging is a classification algorithm that combines multiple versions of the same model. This method uses majority voting for categorical predictions, while for numerical data, it averages these versions of the same model. These multiple versions are constructed using learning sets generated by random sampling from the training data (Breiman, 1996).

$$IG(S, A) = H(S) - \sum \frac{|S_v|}{|S|} H(S_V) \quad (4)$$

Here, S denotes the dataset, $H(S)$ represents the entropy, and S_v corresponds to the subset of S that takes the value v.

$$SplitInfo\ (A) = -\sum \frac{|S_v|}{|S|} log_2 \frac{|S_v|}{|S|} \quad (5)$$

$$\text{J48}=\arg max \frac{IG(S,A)}{SplitInfo(A)} \quad (6)$$

Random Forest

Random Forest is an ensemble learning method that combines multiple decision trees to increase classification accuracy and minimize overfitting. Each tree is trained using a random subset of features, and the classification result is determined by a majority vote among the predictions of all the trees (Belgiu & Dragut, 2016). This method is popularly preferred for large and complex data because it can handle overfitting (Breiman, 2001). The Random Forest method is successful due to its ability to independently construct multiple decision trees under the umbrella of a strong learner (Sun et al., 2018).

$$\hat{y} = mode\{h_1(x), h_2(x), \dots, h_B(x)\} \quad (7)$$

Here, $h_b(x)$ represents the b-th tree.

Performance Evaluation

In this study, statistical metrics such as accuracy, precision, and F-measure were used to understand how well classification models provide reliable and balanced predictions. A confusion matrix is used to analyze the performance of a classification model in detail. Analyzes were carried out using WEKA 3.6 program using machine learning systems (Hall et al., 2009). The confusion matrix is shown in Table 3.

In this context, FP represents False Positive, while FN means False Negative. TN stands for True Negative, while TP indicates True Positive. An evaluation is made by measuring the performance of the classification models using equations 8-15.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FN+FP)} \quad (8)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (9)$$

$$F - Measure = \frac{2\ x\ (Precision\ x\ Recall)}{(Precision+Recall)} \quad (10)$$

$$Recall = \frac{TP}{TP+FN} \quad (11)$$

$$TPR = \frac{TP}{TP+FN} \quad (12)$$

$$FPR = \frac{FP}{FP+TN} \quad (13)$$

$$ROC\ Area = \sum_{i=1}^{n-1}(FPR_{i+1} - FPR_i).\frac{TPR_{i+1}+TPR_i}{2} \quad (14)$$

$$PRC\ Area = \sum_{i=1}^{n-1}(Recall_{i+1} - Recall_i).\frac{Precision_{i+1}+Precision_i}{2} \quad (15)$$

Results and Discussion

This study examines the impact of environments with varying salinity levels on seed germination using T10, T50, T90, T10-90, MGT, and Gmax parameters. Based on these germination characteristics, salinity environments were classified using five machine learning methods, such as Bayes Net, Decision Table, Bagging, J48, and the Random Forest algorithm. A 10-fold cross-validation technique was used to reliably assess the performance and generalization ability of the classification models. The results of the study showed that the parameters in the newly created dataset exhibited very high performance in distinguishing salinity levels. Random Forest, Bagging, and BayesNet classification models successfully classified the data, achieving higher accuracy compared to other models. Table 4 shows the performance metrics of all models used to classify the Corn Genotype dataset.

As seen in Table 4, all classification methods applied to the dataset achieved an accuracy rate above 88.30%. The Random Forest algorithm performed better on this dataset than the other algorithms, achieving 92.02% accuracy, 0.986 roc area, 0.977 prc area, 0.920 precision, and an f-measure metric, making it the most successful model. The J48 model exhibited the lowest classification performance on the dataset, with an accuracy of 88.32%, a precision of 0.882, f-measure of 0.883, a roc area of 0.926, and a prc area of 0.850.

In the process of classifying the Corn Genotype dataset based on salinity, a Correlation-based feature selection (CFS) method was applied to further improve classification performance by removing irrelevant and unnecessary attributes from the data. After applying CFS, genotype and T10-90 attributes were removed from the dataset. The classification results obtained after applying the CFS are shown in Table 5.

When Table 5 is examined, after applying the feature selection method, the most successful classification model in seed germination in different salinity environments was Bayes Net with an accuracy rate of 89.74% (a precision of 0.898, f-measure of 0.898, roc area of 0.975, and prc area of 0.944).

The second most successful model after the Bayes Net model was the Decision Table with an accuracy rate of 89.17% (a precision of 0.892, f-measure of 0.892, roc area of 0.948, and prc area of 0.887).

Table 3. Confusion Matrix.

	Predicted No (0)	Predicted Yes (1)
Actual No (0)	TN	FP
Actual Yes (1)	FN	TP

Table 4. The Performance of all classification models applied to the Corn Genotype dataset.

	Accuracy (%)	Precision	F-Measure	ROC Area	PRC Area
BayesNet	89.74	0.898	0.898	0.971	0.935
Decision Table	88.89	0.889	0.889	0.948	0.886
Bagging	89.17	0.891	0.891	0.975	0.955
J48	88.32	0.882	0.883	0.926	0.850
Random Forest	92.02	0.920	0.920	0.986	0.977

Table 5. The Performances of all classification methods after the feature selection applied to the Corn Genotype dataset

	Accuracy (%)	Precision	F-Measure	ROC Area	PRC Area
BayesNet	89.74	0.898	0.898	0.975	0.944
Decision Table	89.17	0.892	0.892	0.948	0.887
Bagging	88.03	0.880	0.880	0.965	0.925
J48	88.60	0.887	0.886	0.928	0.840
Random Forest	87.46	0.875	0.875	0.968	0.936

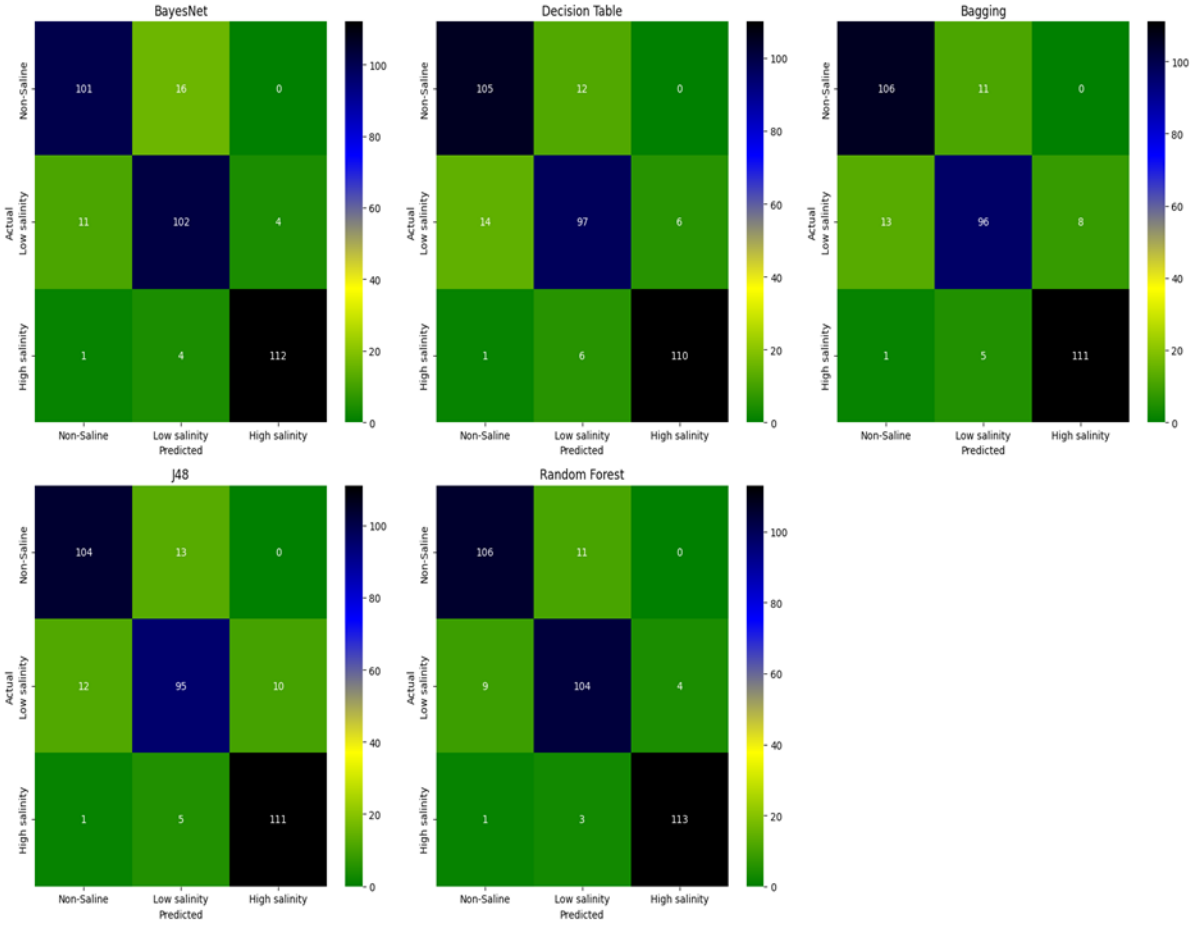

Figure 2. Confusion matrix heatmap of the models obtained using all attributes.

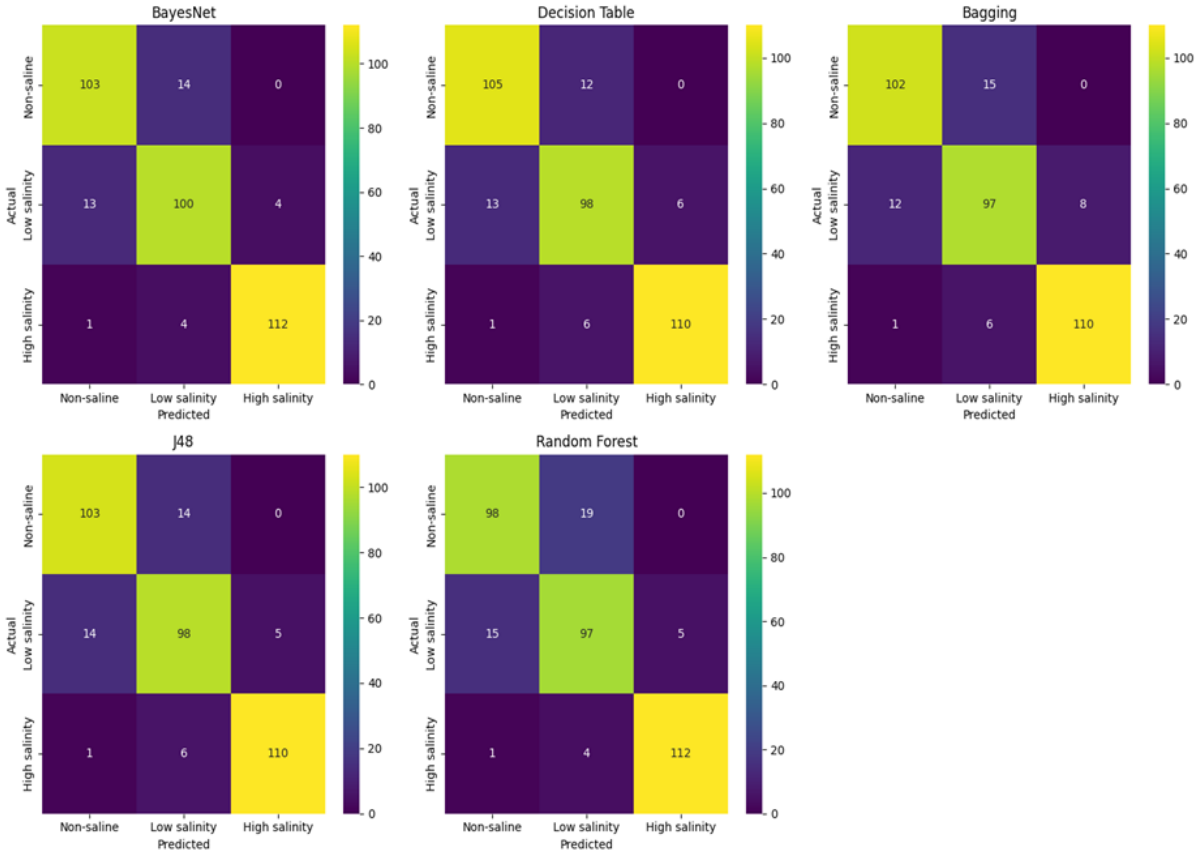

Figure 3. Confusion matrix heat map of the models after applying the CFS

The model with the lowest classification performance was the Random Forest algorithm with an accuracy rate of 87.46%, a precision value of 0.875, an f-measure value of 0.875, a roc area of 0.968, and a prc area of 0.936.

Comparing Tables 4 and 5, the model performance generally improved after removing the T10-90 and genotype attributes from the dataset. A slight decrease in performance was observed in the Random Forest algorithm. In general, it can be said that the models performed better on the dataset when the CFS method was applied, as the data size and complexity were reduced.

In this study, it was observed that after applying the feature selection method, most models showed an improvement in performance, but the Random Forest algorithm experienced a slight decrease in performance. Gandhi and Prabhune (2017) demonstrated that feature selection removes irrelevant and unnecessary features, reduces data size, and potentially improves prediction performance. The fundamental idea is that strategic feature reduction using methods such as CFS can optimize machine learning model performance by focusing on the most informative features. Current literature emphasizes that machine learning methods show similar success in detecting plant stress; Akbari et al. (2024) achieved 98% accuracy in classifying salinity and drought stress in barley using the GenPhen ML approach. Madhumitaa et al. (2024) demonstrated an accuracy rate of 96.79% in soybean stress classification using SVM. These findings were also confirmed in corn, and the performance of the XGBoost model in classification tasks reached 99.3% (Zhang et al., 2025). The research consistently demonstrates that machine learning methods offer non-invasive, scalable solutions for germination screening and stress tolerance prediction in different grain products, thereby becoming a promising tool for agricultural research and product development.

Confusion matrices obtained using all features revealed that the models generally demonstrated high success in distinguishing the three classes. Random Forest and BayesNet models achieved the highest accuracy across all classes, with minimal misclassification rates, particularly in the High Salinity class. J48 achieved moderate classification performance, exhibiting slight cross-confusion in certain classes.

Critically, all models showed sensitivity to differences in salinity levels. Thus, germination parameters have been an important factor in detecting salt stress (Madhumitaa et al., 2024). This situation demonstrates that morphological characteristics can be reliably used to classify plants' responses to salt stress and potentially enable more precise agricultural stress management (Ozkan & Ozkan, 2025).

After applying the feature selection method CFS, the confusion matrices obtained with the remaining features revealed that the classification models were successful in distinguishing three different environments. While BayesNet performs particularly well in the High salinity class, Decision Table performs slightly better in the Non-Saline class. The performances of BayesNet and Decision Table are similar in the Low and High salinity classes. Bagging can be said to offer a generally balanced performance. While J48 has similar results with BayesNet in the Non-Saline class, we can say that it is quite balanced in the Low salinity and High salinity classes. Random Forest and BayesNet provide the highest accuracy rate in High salinity. The classification models are robust in the High salinity class, while there is more confusion in the Non-Saline and Low salinity classes. J48 and Bagging appear to be more suitable for balanced prediction across all classes. If high salinity prediction is critical, it can be better to choose Random Forest or BayesNet.

The recommendation to select machine learning algorithms based on specific prediction requirements is well supported by empirical evidence from numerous studies. Wang et al. (2020) comprehensively compared thirteen algorithms and recommended the Random Forest (RF) method, particularly for salinity mapping in arid environments. RF and BayesNet demonstrate strong performance for high salinity sensitivity. Musleh (2024) further confirmed this by showing that RF and J48 achieved exceptional accuracy (0.993) and sensitivity. For more balanced, cross-class predictions, the J48 and Bagging classifiers prove effective, with the latter achieving 0.992 accuracy. However, studies consistently emphasize that no algorithm performs ideally on all datasets, highlighting the importance of careful and context-specific algorithm selection (Wang et al., 2020). These findings support the role of ML and Deep Learning (DL) methods in addressing plant stress in smart agriculture applications, and a recent review emphasizes that these approaches are indispensable for data analysis and stress management (Zarbakhsh et al., 2025).

Conclusion

This study demonstrates that machine learning models can classify salt stress levels with high accuracy on the corn genotype dataset. The study particularly highlights the critical importance of estimating salt stress level without parameters such as T10-90 and genotype. In the study, classification was first performed using five different models using all parameters in the dataset, and accuracies ranging from 88.32% to 92.02% were achieved. The most successful model was Random Forest, with an accuracy rate of 92.02%. The same models were rerun on the dataset using CFS, a feature selection method, and BayesNet achieved the highest accuracy rate of 89.74%. Bagging achieved the lowest accuracy rate of 88.03%. This study emphasizes that different levels of salt stress can be classified with high accuracy using different machine learning models. It also demonstrates that ML models can successfully predict different levels of salt stress without considering important parameters such as T10-90 and genotype. The results obtained will contribute to both physiological research and agricultural applications.

Declarations

Conflict of Interest

We would like state that there are no conflicts among authors.

Authors' contribution statement

Investigation, formal analysis, methodology, and writing the original draft: MTA; Data collection, supervision, conceptualization, review and editing: GZ.

Financial support

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflict of Interest

The authors declare no conflict of interest

References

Akbari, M., Sabouri, H., Sajadi, S. J., Yarahmadi, S., & Ahangar, L. (2024). Classification and prediction of drought and salinity stress tolerance in barley using GenPhenML. Scientific Reports, 14(1), 17420. https://doi.org/10.1038/s41598-024-68392-w

Altaf, M.T., Liaqat, W., Bedir, M., Comertpay, G., Ali, S.A., Aasim, M., Nadeem, M.A. & Baloch, F.S. (2026). Multivariate analysis and machine learning prediction of Sorghum cultivar traits under nitrogen regulation. BMC Plant Biol. https://doi.org/10.1186/s12870-026-08434-9

Atanda, A. S., Olsen, M., Crossa, J., Burgueño, J., Rincent, R., Dzidzienyo, D., Beyene, Y., Gowda, M., Dreher, K., M. Boddupalli, P., Tongoona, P., Yirenkyi Danquah, E., Olaoye, G., & R. Robbins, K. (2021). Scalable sparse testing genomic selection strategy for early yield testing stage. Frontiers in Plant Science, 12, 658978. https://doi.org/10.3389/fpls.2021.658978

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS journal of photogrammetry and remote sensing, 114, 24 https://doi.org/10.1016/j.isprsjprs.2016.01.011

Botero-Valencia, J., García-Pineda, V., Valencia-Arias, A., Valencia, J., Reyes-Vera, E., Mejia-Herrera, M., & Hernández-García, R. (2025). Machine learning in sustainable agriculture: systematic review and research perspectives. Agriculture, 15(4), 377. https://doi.org/10.3390/agriculture15040377

Breiman, L. (1996). Bagging predictors. Machine learning, 24(2), 123-140.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Carpici, E. B., Celik, N., & Bayram, G. (2009). Effects of salt stress on germination of some maize (*Zea mays* L.) cultivars. African Journal of Biotechnology, 8(19). https://doi.org/10.4314/ajb.v8i19.65187

Chintakunta, A. N., Koganti, S., Nuthakki, Y., & Kolluru, V. K. (2023). Deep learning and sustainability in agriculture: A systematic review. International Journal of Computer Science and Mobile Computing, 12(8), 10-47760. https://doi.org/10.47760/ijcsmc.2023.v12i08.017

Cihan, P., & Kalipsiz, O. (2016). Determining the most successful classification algorithm for the student project questionnaire. TBV J Comput Sci Eng, 8(1), 41-49.

Collado, B. M., Beatriz Aulicino, M., Jacinto Arturi, M., & del Carmen Molina, M. (2016). Selection of maize genotypes with tolerance to osmotic stress associated with salinity. Agricultural Sciences, 7, 82-92. http://dx.doi.org/10.4236/as.2016.72008

Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. Intelligence, 43, 52-64. https://doi.org/10.1016/j.intell.2014.01.004

Gandhi, S. S., & Prabhune, S. S. (2017). Overview of feature subset selection algorithm for high dimensional data. In 2017 International Conference on Inventive Systems and Control (ICISC) (pp. 1-6). https://doi.org/10.1109/ICISC.2017.8068599

Gitmez, M., Karabatak, M., & Varol, A. (2019). Comparison of classification algorithms on optical burst switching network data. In 2019 1st International Informatics and Software Engineering Conference (UBMYK) (pp. 1-4). https://doi.org/10.1109/ubmyk48245.2019.8965474

Guo, S., Lv, L., Zhao, Y., Wang, J., Lu, X., Zhang, M., Wang, R., Zhang, Y., & Guo, X. (2023). Using high-throughput phenotyping analysis to decipher the phenotypic components and genetic architecture of maize seedling salt tolerance. Genes, 14(9), 1771. https://doi.org/10.3390/genes14091771

Ibrar, D., Khan, S., Raza, M., Nawaz, M., Hasnain, Z., Kashif, M., Rais, A., Gul, S., Ahmad, R., & Z. Gaafar, A. R. (2024). Application of machine learning for identification of heterotic groups in sunflower through combined approach of phenotyping, genotyping and protein profiling. Scientific Reports, 14(1), 7333. https://doi.org/10.1038/s41598-024-58049-z

Isik, H., Tasdemir, S., Selim Taspinar, Y., Kursun, R., Cinar, I., Yasar, A., Tahsin Yasin, E., & Koklu, M. (2023). Maize seeds forecasting with hybrid directional and bi-directional long short-term memory models. Food Science & Nutrition, 12(2), 786-803. https://doi.org/10.1002/fsn3.3783

ISTA. (2015). International Rules for Seed Testing. International Seed Testing Association.

Keleş, A. E., & Keleş, M. K. (2018). Prediction of employee leadership perception in construction management using feature selection with genetic algorithm. Journal of Çukurova University Faculty of Engineering and Architecture, 33(4), 97-110. https://doi.org/10.21605/cukurovaummfd.525060

Khalid, N., Tarnawa, Á., Balla, I., Omar, S., Abd Ghani, R., Jolánkai, M., & Kende, Z. (2023). Combination effect of temperature and salinity stress on germination of different maize (*Zea mays* L.) varieties. Agriculture, 13(10), 1932. https://doi.org/10.3390/agriculture13101932

Khan MA, Ahmed MZ, Hameed A. (2006). Effect of sea salt and L-Ascorbic acid on the seed germination of halophytes. Journal of Arid Environments, 67 (3): 535-540.

Luo, M., Zhao, Y., Weisong, W., Zhang, R., Su, A., Li, C., Wang, X., Xing, J., Shi, Z., & Zhao, J. (2017). Effect of saline stress on the physiology and growth of maize hybrids and their related inbred lines. Maydica, 62(2), 8.

Madhumitaa, P. S., Ragavi, C., Kiranmayi, C., & Prabhavathy, P. (2024). Drought and salinity stress classification in soyabean crops: Comparative analysis of machine learning models. In 2024 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT) (pp.1-5). https://doi.org/10.1109/IConSCEPT61884.2024.10627854

Maheswari, M. U., & Ramani, R. (2023). A comparative study of agricultural crop yield prediction using machine learning techniques. In 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1428-1433). https://doi.org/10.1109/ICACCS57279.2023.10112854

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. ACM SIGKDD explorations newsletter, 11(1), 10-18. https://doi.org/10.1145/1656274.1656278

Mohanavelu, A., Naganna, S. R., & Al-Ansari, N. (2021). Irrigation induced salinity and sodicity hazards on soil and groundwater: An overview of its causes, impacts and mitigation strategies. Agriculture, 11(10), 983. https://doi.org/10.3390/agriculture11100983

Musleh, F. A. (2024). A comprehensive comparative study of machine learning algorithms for water potability classification. Int. J. Com. Dig. Sys, 15(1). https://doi.org/10.12785/ijcds/150184

Ozkan, S. S., & Ozkan, E. C. (2025). The effect of salinity stress on the germination and early growth parameters of selected ryegrass species. ISPEC Journal of Agricultural Sciences, 9(1), 36-47. https://doi.org/10.5281/zenodo.14160787

Pitman, M. G., & Läuchli, A. (2002). Global impact of salinity and agricultural ecosystems. Salinity: environment-plants-molecules, 3, 20. https://doi.org/10.1007/0-306-48155-3_1

Premasudha, B. G., Thara, D. K., & Tara, K. N. (2022). Ml based methods xgboost and random forest for crop and fertilizer prediction. In 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 492-497). IEEE. https://doi.org/10.1109/CICN56167.2022.10008234

Sun, Z., Sun, H., & Zhang, J. (2018). Multistep wind speed and wind power prediction based on a predictive deep belief network and an optimized random forest. Mathematical Problems in Engineering, 2018(1), 6231745. https://doi.org/10.1155/2018/6231745

Tong, H., & Nikoloski, Z. (2021). Machine learning approaches for crop improvement: Leveraging phenotypic and genotypic big data. Journal of plant physiology, 257, 153354. https://doi.org/10.1016/j.jplph.2020.153354

Wang, F., Shi, Z., Biswas, A., Yang, S., & Ding, J. (2020). Multi-algorithm comparison for predicting soil salinity. Geoderma, 365, 114211. https://doi.org/10.1016/j.geoderma.2020.114211

Zarbakhsh, S., Fakhrzad, F., Rajkovic, D., Niedbała, G., & Piekutowska, M. (2025). Approaches and challenges in machine learning for monitoring agricultural products and predicting plant physiological responses to biotic and abiotic stresses. Current Plant Biology, 100535. https://doi.org/10.1016/j.cpb.2025.100535

Zhang, L., Zhang, F., Du, W., Hu, M., Hao, Y., Ding, S., ... & Zhang, D. (2025). Machine learning analysis of maize seedling traits under drought stress. Biology, 14(7), 787. https://doi.org/10.3390/biology14070787

Zulkadir, G. (2025). Salt tolerance of different maize genotypes during germination and seedling stages. Phyton, 94(6), 1879. https://doi.org/10.32604/phyton.2025.064144