

EJONS

International Journal on Mathematic, Engineering and Natural Sciences

(Uluslararası Fen, Mühendislik ve Doğa Bilimleri Dergisi)

<https://ejons.org/index.php/ejons>

e-ISSN: 2602 - 4136

Review Article

Doi: <https://doi.org/10.5281/zenodo.14253789>

Baklagil Islahında Makine Öğrenimi: Genotip ve Çevre Etkileşimlerinin Modellenmesi

Sibel IPEKESEN¹, Süreyya Betül RUFATIOĞLU², *, Murat TUNC³, Behiye Tuba BICER¹¹Dicle University, Agriculture Faculty, Department of Field Crops, Diyarbakir, Türkiye.²Harran University, Agriculture Faculty, Department of Soil Science and Plant Nutrition, Sanliurfa, Türkiye.³Harran University, Agriculture Faculty, Department of Field Crops, Sanliurfa, Türkiye.*Corresponding Author: sureyyarufaoglu@harran.edu.tr

Makale Tarihi

Geliş: 24.09.2024

Kabul: 13.11.2024

Anahtar Kelimeler

Makine öğrenmesi,
Baklagiller,
Bitki ıslahı,
Yapay sinir ağları,
ÇevreGenotip
interaksiyonu.

Öz: Bu derlemede, baklagil ıslahında genotip-çevre (G×E) etkileşimlerinin modellenmesinde makine öğrenimi tekniklerinin önemi ele alınmaktadır. Tarımsal üretim, iklim değişikliği ve çevresel stres faktörlerinden büyük ölçüde etkilenmektedir ve bu etkileşimlerin daha iyi anlaşılması, çevresel olarak uyumlu ve yüksek verimli çeşitlerin geliştirilmesi için kritik öneme sahiptir. Genotip-çevre etkileşimleri, genetik ve çevresel faktörlerin bitki performansını nasıl etkilediğini anlamak için kritik öneme sahiptir. Çalışmada, destek vektör makineleri (SVM), rastgele ormanlar (RF), derin öğrenme (DL) ve sinir ağları (ANN) gibi makine öğrenimi algoritmalarının yüksek boyutlu veri setlerini işlemek, genotip stabilitesini değerlendirmek ve çevresel streslere tepkiyi modellemek için kullanıldığı bildirilmiştir. Bu teknikler baklagillerde verim, kalite ve adaptasyon gibi karmaşık özelliklerin değerlendirilmesinde önemli avantajlar sunmaktadır. Makine öğrenimi yaklaşımları, geleneksel yöntemlerin eksikliklerinin üstesinden gelmekte ve özellikle büyük veri setlerinin işlenmesinde, çevresel değişkenlerin modellenmesinde ve genetik performansın tahmin edilmesinde daha güçlü araçlar sağlamaktadır. Bu yöntemlerin etkin kullanımı, tarımsal üretimde sürdürülebilirliği artırmak ve iklim değişikliğine dayanıklı çeşitler geliştirmek için büyük bir potansiyele sahiptir.

Atf Künyesi: Ipekesen, S., Rufatioğlu, S. B., Tunç, M., Bicer, B.T. (2024). Baklagil Islahında Makine Öğrenimi: Genotip ve Çevre Etkileşimlerinin Modellenmesi, EJONS International Journal on Mathematic, Engineering and Natural 8 (4) 493-503. **How To Cite** Ipekesen, S., Rufatioğlu, S. B., Tunç, M., Bicer, B.T. (2024). Machine Learning in Legume Breeding: Modelling Genotype and Environment Interactions EJONS International Journal on Mathematic, Engineering and Natural 8 (4) 493-503.

Machine Learning in Legume Breeding: Modelling Genotype and Environment Interactions

Article Info

Received: 24.09.2024

Accepted: 13.11.2024

Abstract: In this review, addresses the importance of machine learning techniques in modelling genotype-environment (G×E) interactions in legume breeding. Agricultural production is greatly affected by climate change and environmental stressors, and a better understanding of these interactions is critical for the development of environmentally adaptive and high-yielding varieties. Genotype-

Keywords

Machine learning
 Legumes,
 Plant breeding,
 Artificial neural network,
 Genotypeenvironment
 interactions

environmental interactions are valuable for understanding how genetic and environmental factors affect plant performance. The study reported that machine learning algorithms such as support vector machines (SVM), random forests (RF), deep learning (DL) and neural networks (ANN) are used for processing high-dimensional data sets, assessing genotype stability and modelling response to environmental stresses. These techniques offer significant advantages in the evaluation of complex traits such as yield, quality and adaptability in legumes. Machine learning approaches overcome the shortcomings of traditional methods and provide more powerful tools, especially in processing large data sets, modelling environmental variables and predicting genetic performance. The effective use of these methods has great potential to increase sustainability in agricultural production and to develop varieties that are resilient to climate change.

1. Introduction

Agricultural production has faced many challenges with climate change. Climate change has many direct and indirect negative effects on agricultural production, such as loss of yield, changes in plant development and harvest timing, reduction of water resources, decrease in soil quality, increase in disease and pest populations, and loss of nutritional value of plant products. Extreme temperatures, droughts, floods and other extreme weather events accelerate agricultural plants' growth and maturation processes, reducing yield and product quality. Excessive rainfall and erosion reduce soil quality, leading to a decrease in suitable land for agriculture, while production is restricted in water-dependent agricultural areas. In addition, it increases the spread of plant diseases and pests, requiring the use of chemicals, thus negatively affecting environmental health. Summarily, climate change threatens sustainability in agriculture and poses serious risks to food security, economic income and rural life. Therefore, minimizing the difficulties of agricultural output to meet the world population's nutritional needs, which are expected to exceed nine billion by 2050, will increase the sustainability of production. The importance of solutions such as the development of resistant varieties, precision agriculture and climate-friendly practices is increasing.

However, current yield trends in agricultural production are insufficient to achieve these goals (Keating et al., 2014). Therefore, researchers need to shorten the breeding process and adopt new strategies by utilizing approaches such as next-generation sequencing (NGS), advanced phenotyping platforms, and machine learning (ML) to increase production (Drovandi et al., 2017). The development of next-generation strategies in plant breeding is critical for leading a new revolution. In addition to traditional methods and biotechnological applications used in plant breeding, machine learning strategies have recently played an important role in rapidly generating, evaluating and searching for plant-appropriate models of complex large data sets. A Flowchart for using machine-learning algorithms in agricultural production was given in Figure 1.

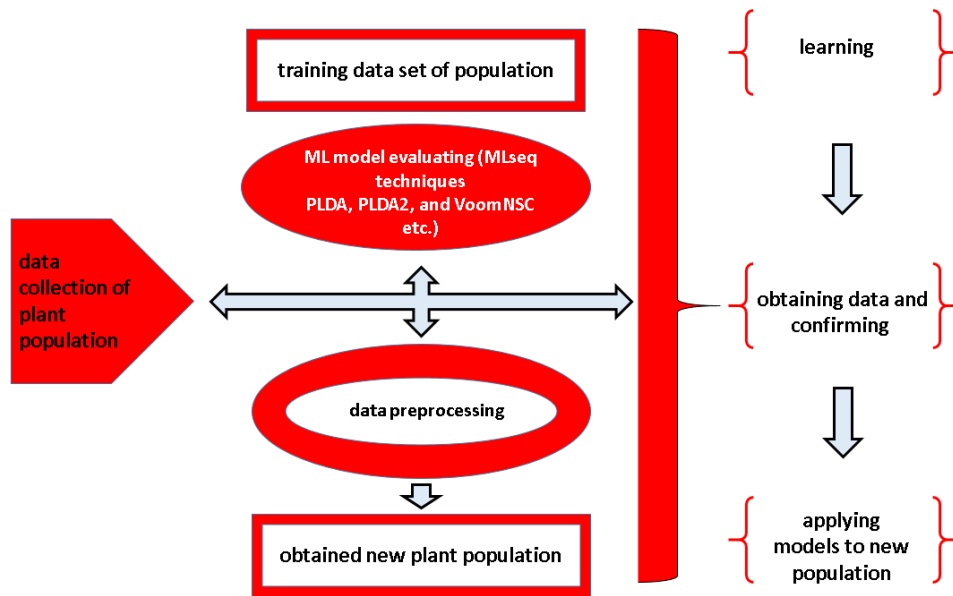


Figure 1: Presentation of a flowchart for using machine-learning algorithms in agricultural production

Machine learning techniques, which can be used as input for algorithms in developing and training accurate prediction models, will have an important place in accelerating the development of resistant crops by identifying important relationships that regulate a biological process (Cheng and Wang, 2024). Machine learning models can be applied to cereals, legume crops, oil crops etc. for yield prediction (Figure 2).

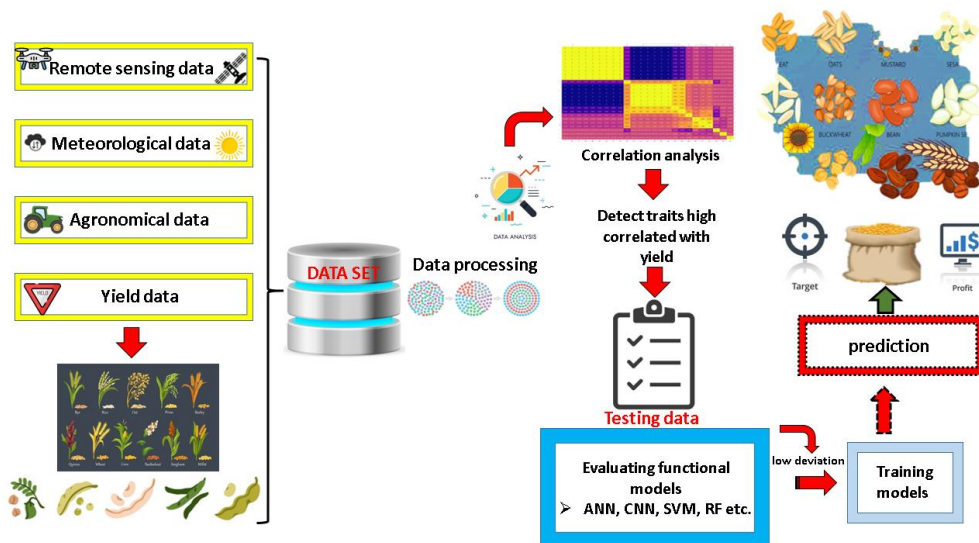


Figure 2: Machine learning models applying yield prediction of cereals, legume crops, oil crops etc.

The review purposes to discuss the importance of ML techniques in plant breeding and their usability in the breeding of legume plants. We provide information about the benefits of sensing technologies and phenotyping platforms in agriculture, the effective use of ML algorithms in large data sets, and studies conducted on legume crop breeding based on machine learning.

2. Using machine learning in plant breeding

Since yield in plants is a complex quantitative trait controlled by several genes and strongly influenced by the environment, improving these complex traits by direct selection is impossible. In legume plants, genotype-environment (G×E) interaction describes the interaction between the genetic structure (genotype) of the plant and the environment in which it is grown (factors such as climate, soil, and maintenance practices). G×E interaction is critical for understanding how genetic and environmental factors affect plant performance. This interaction is essential for developing more efficient and adaptive genotypes in plant breeding.

This interaction is particularly important in the evaluation of traits such as yield, quality, and adaptation. Since legumes can be grown in a wide range of environmental conditions, the effects of genotype-environment interaction are quite pronounced. Genotype-environment interactions occur when different genotypes perform differently under changing environmental conditions. This can manifest itself as a particular genotype performing well in one environment but poorly in another. Many factors such as temperature, precipitation, light intensity, day length, soil pH, salinity, organic matter content and nutrient amount, fertilization, irrigation, planting time, planting density, biotic and resistance levels of genotypes to biotic factors affect the G×E interaction (Figure 3). G×E interaction is very important to understand whether legume genotypes are adapted to a certain region (adaptation), to understand the yield potential of genotypes in different environmental conditions and for breeding studies.

Therefore, an ideal genotype should perform stable high adaptive according to environmental conditions. AMMI (additive main effects and multiplicative interaction), GGE biplot analysis and stability analyses are performed to evaluate the G×E interaction. Plant breeders interpret the relationships between traits with univariate and multivariate linear methods such as correlation coefficient analysis, PCA and multiple regression analysis (Ipekesen et al., 2023). Environmental changes and genotype-environment interactions cause differences in yield and phenotypic traits of varieties.

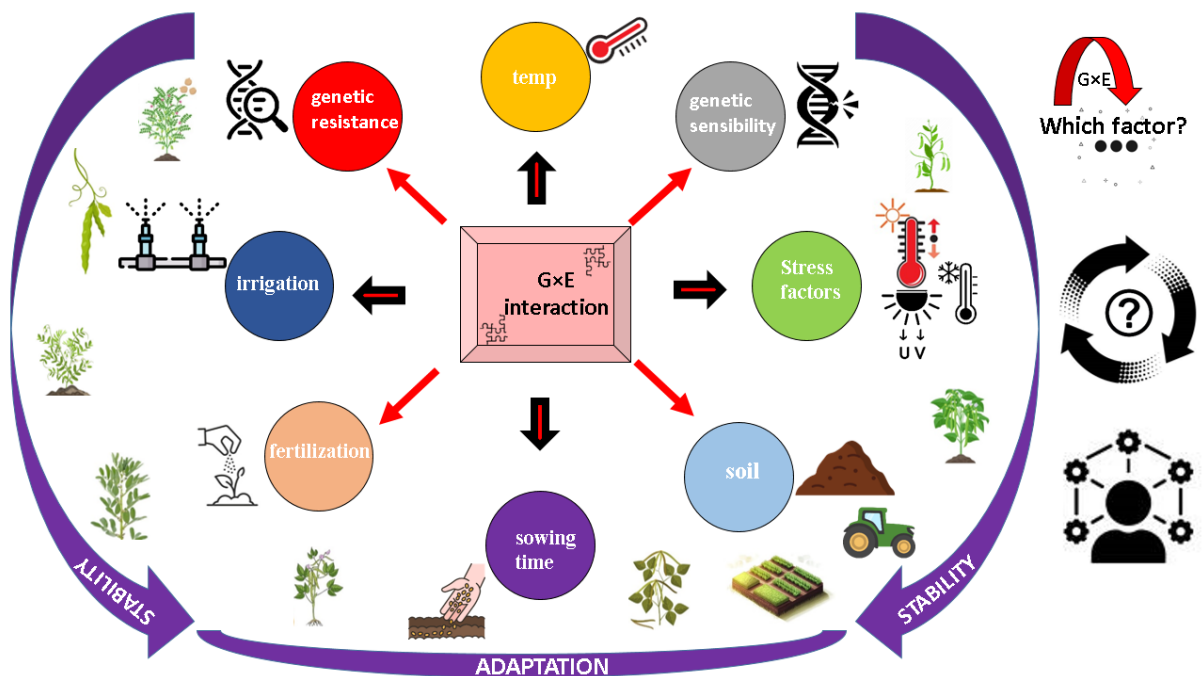


Figure 3: Factors affecting G×E interaction

Stability analysis can be used effectively to evaluate the relative performance of genotypes in different environments (Ajay et al., 2020; Ipekesen et al., 2023). In addition, many classical univariate

approaches such as regression analysis (Finlay and Wilkinson, 1963), coefficient regression (Eberhart and Russell, 1966), equivalence analysis (Wricke, 1962), stability variance analysis (Shukla, 1972) and coefficient variance analysis (Francis et al., 1978) are used in the interpretation of genotype-environment interaction. However, the distribution and homogeneity of variance must be tested for data before using these parametric methods (Flores et al., 1988). In addition, GXE interaction can be explained by taking a space view of a genotype with cluster analysis, multiplicative interactions (AMMI) and genotype-environment interaction (GGE biplot) (Erdemci, 2018; Ipekesen et al., 2023). However, all these methods are multivariate stability analyses, and these methods are complex methods, which can not provide a simple measure of the yield stability of genotypes (Flores et al., 1978). Recently, the accuracy of the studies has been strengthened by applying linear and nonlinear regression models together (Li et al., 2020).

G×E interaction is an important component to be considered to increase the success of breeding programs. Traditional methods use statistical models to analyze the performance of genotypes under environmental conditions (e.g. AMMI or GGE biplot analyses). However, these methods often have limitations in handling high-dimensional datasets or modelling environmental variability in detail. Data sources of G×E interactions are usually large and complex. Machine learning provides a powerful tool to overcome these challenges in agricultural production (Moosavi ve Sepaskha 2012). Machine learning methods can be listed below.

3. Machine Learning Algorithms

3.1. Supervised Learning Methods: Many researchers are focused on supervised learning such as RM, CNN, RNN and SVM (Gonzalez et al., 2018).

3.1.1. Regression models (RM): Used to predict genotype performance based on environmental factors and genetic characteristics.

3.1.2. Support vector machines (SVM): Used to classify G×E interactions and make performance predictions.

3.1.3. Deep learning (DL): Neural networks (CNN and RNN) are used to learn complex relationships from large genotype and environment data sets (Kotsiantis et al. 2006).

3.2. Unsupervised Learning Methods

3.2.1. Clustering algorithms: Used to group genotypes or environments based on environmental and genotypic similarities (K-means and hierarchical clustering).

3.2.2. Dimensionality reduction: Algorithms such as PCA or t-SNE are used to identify important environmental and genetic factors.

3.3. Ensemble Methods

3.3.1. Random Forest (RF): Used to identify important factors affecting G×E interactions and to make yield predictions.

3.3.2. Gradient Boosting (GB): Especially effective for increasing the predictive power when working with multidimensional and complex data sets. However, supervised and unsupervised machine learning algorithms are commonly used compared to other algorithms.

Machine learning can identify complex relationships between genetic data and environmental variables in the estimation of G×E interactions using deep learning (DL) algorithms. The performance stability of genotypes across years and locations can be assessed with machine learning (ML) models. The adaptation of genotypes to specific environments can also be estimated with ensemble models

and regression algorithms. ML algorithms can be used, especially for data cleaning in large and multidimensional data sets, dimensionality reduction (e.g. PCA) and feature selection. SNP (Single Nucleotide Polymorphism) data file provides information about genetics. Environmental data provides information about temperature, precipitation, soil properties, and other agro climatic factors. Performance data give information about measurements such as yield, quality and stress tolerance.

ML algorithms have been successfully applied to big data analytics, environmental prediction, genetic variation analysis, and genotype stability assessment (Akmel et al., 2018). Thanks to modern breeding techniques and statistical analysis, the response of genotypes to environmental variations is better understood and it is possible to develop more resistant, high-yielding varieties.

Machine learning can make revolutionary changes in plant breeding, increasing the efficiency and speed of traditional methods (Soares et al., 2013). In classical plant breeding studies, it is quite difficult to analyse the nonlinear relationships of dependent and independent variables with regression-based models. ML techniques have an important role in accelerating the development of resistant species by identifying significant relationships, which regulate biological processes. The flowchart for using machine-learning algorithms in plant breeding was given in Figure 1.

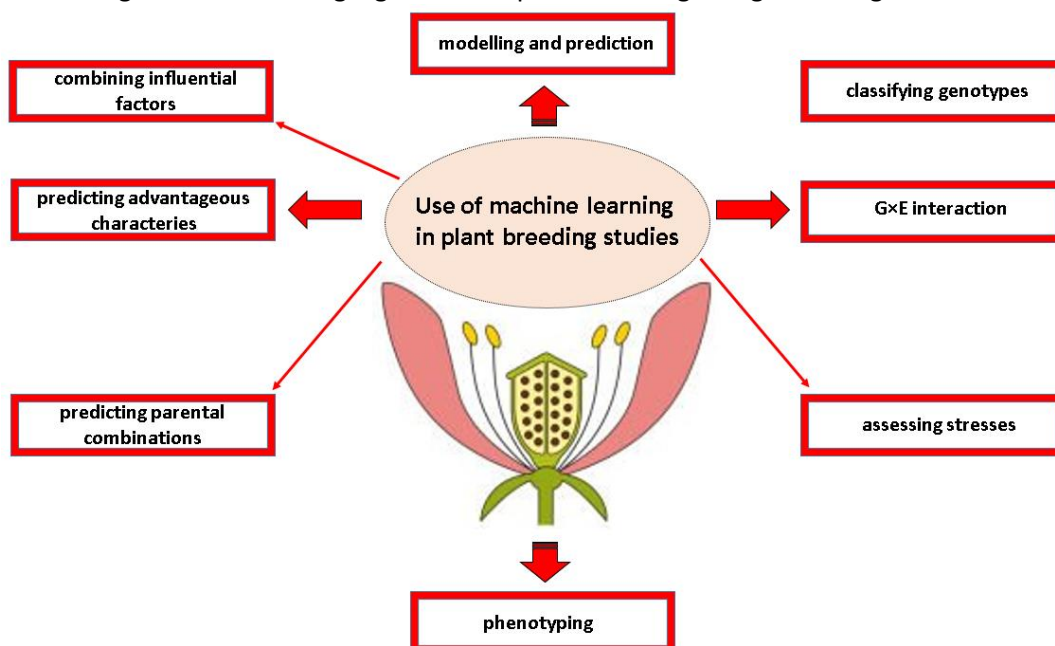


Figure 4: Presentation of a flowchart for using machine-learning algorithms in plant breeding

Several researchers have investigated the shape, area, size and colour characteristics of legume seeds using traditional methods (Jogihalli et al. 2017; Sastry et al. 2019). Since these physical measurements are time-consuming and laborious, new technologies have been developed to overcome these problems. Image processing techniques and spectrophotometric analysis, as well as discriminant analysis, principal component analysis and hierarchical cluster analysis methods, facilitate accurate and effective selection and grouping of varieties. Cetin et al., (2023) reported that the physical properties of 26 different chickpea varieties were better than other machine learning algorithms in Random Forest (RF) and Multilayer Perceptron (MP) machine learning models and that the results obtained were important for selection and breeding studies. They also added that computer and machine vision techniques could provide effective and practical tools for agricultural machinery designers and seed companies by providing accurate and rapid classification of chickpea varieties. Additionally, Kirtis et al., (2022) stated that the RF model gave the best performance with high accuracy among vector regression, Gaussian process regression, XGBoost, random forest (RF) models and multilayer perceptron neural network machine learning algorithms in predicting shoot number and length in chickpea.

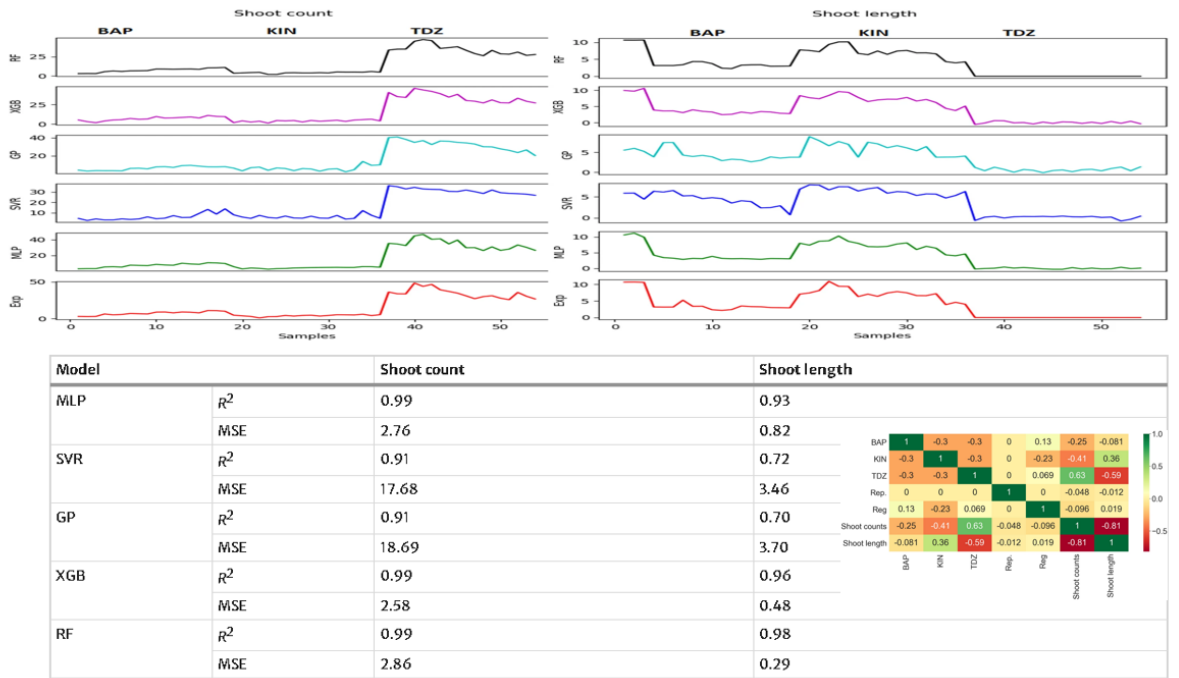


Figure 5: Machine learning models can be used in shoot count and shout length of chickpea varieties (reported by Kirtis et al., 2022).

It is possible to apply machine learning algorithms such as ANN and MLR to estimate the nonlinear relationships between dependent and independent variables, the contribution of yield components to yield, and quantitative traits (Emamgholizadeh et al., 2015). ANN also enables the use of image-based CNN, DNN and ARIMA models to detect the relationship between meteorological data and mineral fertilizer use with yield (Niedbała et al., 2019; Sadeghi-Tehran et al., 2019; Haider et al., 2019). DL algorithms such as CNN, k-nearest neighbour, probabilistic neural network, support vector machine and genetic algorithm can be used to determine the genetic diversity of plant genotypes and classify morphological parameters (Pandolfi et al., 2009; Saini et al., 2020). Yan et al. (2019) also reported that CNN and SVM algorithms can be used to select between morphologically similar species with an accuracy of 96.7% and 74.6%, respectively, and Azizi et al. (2016) reported that the ANN model gave successful results with an accuracy of 100% in the classification of morphological features and varieties through image processing.

Analyzing large data sets of plant genetic material enables understanding genotype-phenotype relationships, identifying genes controlling agronomic traits, and optimizing genomic selection strategies (Schrider and Kern, 2018). Efficient genetic variant estimation can be achieved with algorithms including support vector machines (SVM) or RF (Cortes and Vapnik, 1995; Gaonkara and Davatzikos, 2013). Additionally, Ayele (2020) stated that the decision tree (Ctree) was the best performing algorithm in terms of sensitivity and specificity in classifying chickpea varieties with an accuracy rate of 97.5%.

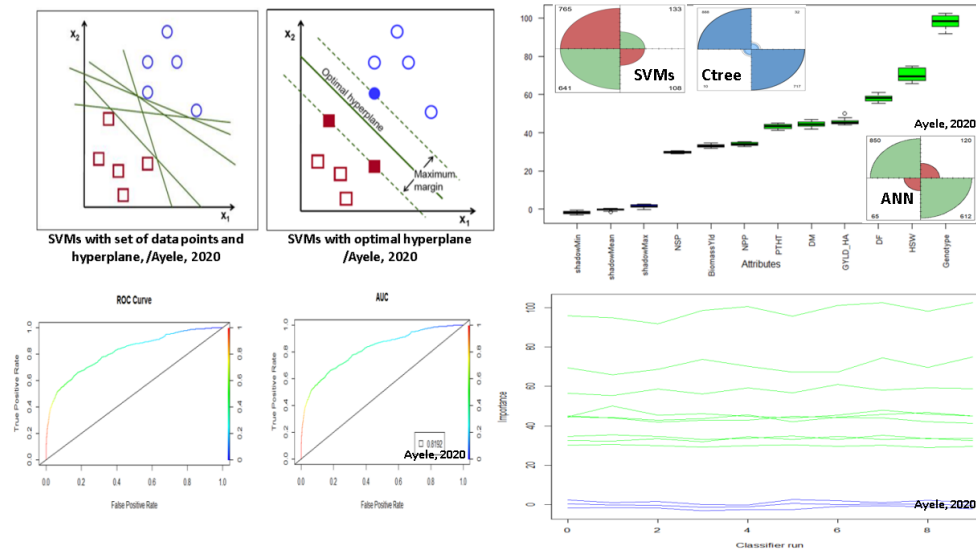


Figure 6: Use of machine learning approaches for chickpea phenotyping

The phenotyping process, which defines the measurement of the physical and biological characteristics of the plant, requires a lot of labour and time. Thanks to the image processing techniques used in machine learning algorithms, features such as plant growth, leaf area, and disease symptoms can be measured quickly. In addition, images collected with drones and thermal cameras can be easily processed and analysed using the CNN model (Lee et al., 2018). Therefore, the phenotyping process, which takes a lot of time with classical breeding, can be shortly completed. Additionally, these models can be used in the early diagnosis of plant diseases, which play an important role in maintaining agricultural productivity. Disease symptoms can be identified by imaging data of ML algorithms and the spread dynamics of the disease can be modelled. Moreover, machine learning models such as time series analysis and spatial data processing can also be used to learn how plants respond to environmental stress (drought, heat, salinity, etc.).

Machine learning can analyze large data sets with multi-disciplinary agricultural technologies and high-performance computing in agricultural production, data processing, measurement and evaluation models (Liakos et al., 2018). In addition, these techniques can be used to determine the correlation between plant performance, environmental response and gene function and to determine physiological changes in plants caused by stress thanks to advanced phenotyping technologies such as sensing technologies (Triposodi et al., 2018). These models facilitate adaptation studies by enabling the development of stress-resistant plant species. Additionally, GBLUP and Bayesian model facilities selection from plants with the best genetic combinations for the study, predicting high accuracy from genetic data. Esposito et al., (2018) stated that ML could quickly collect big data sets of genomics and phenomics to determine the relationships between traits and alleles, and could be produced appropriate models. Yan and Wang (2023) stated that ML algorithms could determine the evaluation of data size, inference of gene regulatory networks, finding of genes, and correlation relationship between genotypes and phenotypes. Parmly et al., (2019) reported that the RF model gave high-accuracy results in predicting the change in seed yield at different row spacing and planting density by utilizing the data of high-dimensional phenotypic traits. Yoosefzadeh Najafabadi et al. (2023) stated that high-throughput omics technologies, machine learning methods, and bioinformatics tools can accurately analyse large and complex datasets and can be used to develop high-yielding, climate change, pest, and disease-resistant varieties. Machine learning models are effectively used for yield prediction with priority indicators. Zhao et al., (2016) reported that machine learning approaches RF and SVM can be used for character selection and variety classification based on complex phenotypic datasets. Taghavi Namin et al. (2018) reported that the CNNLSTM model outperformed SVM, LSTM, CNN, CNN+CRF, and CNN-LSTM models by 93% in the phenotyping and classification of plants. Kohzadi et al. (1996) reported that the autoregressive integrated moving average model (ARIMA) model, Zou

et al. (2007) reported feedforward neural network (FFNN) and time series models, Michael D, Johnson et al. (2016) reported that multiple linear regression (MLR), machine learning models Bayesian neural networks (BNN) and model-based recursive partitioning were successful in predicting crop yields. Zhang et al (2021) used phenomic technologies to monitor and predict seed yield, days to 50% flowering, and days to physiological maturity of chickpea and pea grown in three different locations. The study utilized multi-spectral UAV-based imaging and LASSO regression and reported that image-based features provided more predictive results than LASSO regression models. Moreover, Varol et al. (2023) used MLP and RF models for mass, colour estimation of chickpea seeds under varying irrigation water amounts using ML and reported that the MLP model gave better results than these models.

4. Conclusion

Genotype-environment (G×E) interaction is an important research topic in plant genetics. Although traditional statistical methods have an important place in the analysis of G×E interaction, machine learning (ML) approaches have revolutionized this field in recent years. In particular, ML has many advantages in processing large data sets, modelling environmental variables, and predicting the environmental adaptation of genotypes. These methods offer strong potential, especially for understanding the effects of environmental variations on genotypic performance and developing resistant cultivars. Researchers in machine learning have reported many studies. The usability of ML in plant breeding will increase by the collection and analysis of more comprehensive data sets in the future.

To increase the efficiency of the use of machine learning techniques in plant breeding studies.

- There is a need to develop advanced systems for genotyping and phenotyping in plant breeding studies.
- Machine learning-supported models of phenotyping need to be developed in changing environmental conditions.
- It is necessary to increase the applicability of phenotyping and genotyping models in genomic selection-based production programs of legume plants.
- To make the genetic gain and genomic prediction of legume plants accurate and effective, the use of high-density genotyping and phenotyping needs to be expanded.

References

- Ajay, B.C.; Bera, S.K.; Singh, A.L.; Kumar, N.; Gangadhar, K.; Kona, P. Evaluation of Genotype Environment Interaction and Yield Stability Analysis in Peanut Under Phosphorus Stress Condition Using Stability Parameters of AMMI Model. *Agric. Res.* 2020, 1–10.
- Ajay, B.C.; Bera, S.K.; Singh, A.L.; Kumar, N.; Gangadhar, K.; Kona, P. Evaluation of Genotype Environment Interaction and Yield Stability Analysis in Peanut Under Phosphorus Stress Condition Using Stability Parameters of AMMI Model. *Agric. Res.* 2020, 1–10.
- Azizi, A.; Abbaspour-Gilandeh, Y.; Nooshyar, M.; Afkari-Sayah, A. Identifying Potato Varieties Using Machine Vision and Artificial Neural Networks. *Int. J. Food Prop.* 2016, 19, 618–635.
- C. Gonzalez, M. Juan, L. Ornella, P. Perez Rodrigue, D. Gianola, S. Dreisigacker and J. Crossa, "Applications of Machine Learning Methods to Genomic Selection in Breeding Wheat for Rust Resistance," *The Plant Genome*, 2018.
- Çetin, N., Ozaktan, H., Uzun, S., Uzun, O., & Ciftci, C. Y. (2023). Machine learning based mass prediction and discrimination of chickpea (*Cicer arietinum* L.) cultivars. *Euphytica*, 219(1), 20.
- Cheng, Q., & Wang, X. (2024). Machine Learning for AI Breeding in Plants. *Genomics, Proteomics & Bioinformatics*, 22(4), qzae051.
- Cortes, C.; Vapnik, V. Support vector networks. *Mach. Learn.* 1995, 20, 273–297.
- Drovandi, C.C.; Holmes, C.; McGree, J.M.; Mengersen, K.; Richardson, S.; Ryan, E.G. Principles of experimental design for big data analysis. *Stat. Sci.* 2017, 3, 385–404.

- Eberhart, S.A.; Russell, W.A. Stability Parameters for Comparing Varieties 1. *Crop Sci.* 1966, 6, 36–40.
- Emamgholizadeh, S.; Parsaeian, M.; Baradaran, M. Seed yield prediction of sesame using artificial neural network. *Eur. J. Agron.* 2015, 68, 89–96.
- Erdemci, İ. (2018). Investigation of genotypex environment interaction in chickpea genotypes using AMMI and GGE biplot analysis. *Turkish Journal of field crops*, 23(1), 20-26.
- Esposito, S.; Aversano, R.; Bradeen, J.M.; Di Matteo, A.; Villano, C.; Carputo, D. Deep-sequencing of *Solanum commersonii* small RNA libraries reveals riboregulators involved in cold stress response. *Plant Biol.* 2019, in press.
- F. Akmel, E. Birihanu and B. siraj, "A Literature Review Study of Software Defect Prediction using Machine Learning Techniques," *International Journal of Emerging Research in Management and Technology*, vol. 6, pp. 300-306, 2018
- Finlay, K.; Wilkinson, G. The analysis of adaptation in a plant-breeding programme. *Aust. J. Agric. Res.* 1963, 14, 742. [CrossRef] 83. Eberhart, S.A.; Russell, W.A. Stability Parameters for Comparing Varieties 1. *Crop Sci.* 1966, 6, 36–40.
- Flores, F.; Moreno, M.; Cubero, J. A comparison of univariate and multivariate methods to analyze GE interaction. *Field Crop. Res.* 1998, 56, 271–286.
- Francis, C.A.; Prager, M.; Laing, D.R.; Flor, C.A. Genotype Environment Interactions in Bush Bean Cultivars in Monoculture and Associated with Maize. *Crop Sci.* 1978, 18, 237–242.
- G. Tzanis, I. Katakis, I. Partalas and I. Vlahavas, "Modern applications of machine learning," in *Proceedings of the 1st Annual SEERC Doctoral Student Conference DSC*, vol. 1, 2006, pp. 1-10.
- Gaonkar, B.; Davatzikos, C. Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification. *Neuroimage* 2013, 10, 78270–78283.
- Ghaffari, H.; Tadayon, M.R.; Razmjoo, J.; Bahador, M.; Soureshjani, H.K.; Yuan, T. Impact of Jasmonic Acid on Sugar Yield and Physiological Traits of Sugar Beet in Response to Water Deficit Regimes: Using Stepwise Regression Approach. *Russ. J. Plant Physiol.* 2020, 67, 482–493.
- H. Zou, G. Xia, F. Yang and H. Wang, "An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecastin," *Neurocomputing*, vol. 70, pp. 2913-2923, 2007.
- Haider, S.A.; Naqvi, S.R.; Akram, T.; Umar, G.A.; Shahzad, A.; Sial, M.R.; Khaliq, S.; Kamran, M. LSTM Neural Network Based Forecasting Model for Wheat Production in Pakistan. *Agronomy* 2019, 9, 72.
- Ipekesen, S., Akyildiz, M. I., & Alp, A. (2023). Evaluation of yield and quality traits of advanced bread wheat lines under Diyarbakir province irrigated conditions by GGE biplot technique.
- Jogihalli P, Singh L, Kumar K, Sharanagat VS (2017) Physico-functional and antioxidant properties of sand-roasted chickpea (*Cicer arietinum*). *Food Chem* 237:1124–1132
- K. Liakos, P. Busato, D. Moshou, S. Pearson and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, p. 2674, 2018.
- Keating, B.A.; Herrero, M.; Carberry, P.S.; Gardner, J.; Cole, M.B. Food wedges: Framing the global food demand and supply challenge towards 2050. *Glob. Food Secur.* 2014, 3, 125–132.
- Kirtis, A., Aasim, M., & Katırcı, R. (2022). Application of artificial neural network and machine learning algorithms for modeling the in vitro regeneration of chickpea (*Cicer arietinum* L.). *Plant Cell, Tissue and Organ Culture (PCTOC)*, 150(1), 141-152.
- Lee, U.; Chang, S.; Putra, G.A.; Kim, H.; Kim, D.H. An automated, high-throughput plant phenotyping system using machine learning-based plant segmentation and image analysis. *PLoS ONE* 2018, 13, e0196615.
- Li, N.; Lin, H.; Wang, T.; Li, Y.; Liu, Y.; Chen, X.; Hu, X. Impact of climate change on cotton growth and yields in Xinjiang, China. *Field Crop. Res.* 2020, 247, 107590.
- M. D. Johnson, W. W. Hsieh, A. J. Cannon, A. Davidson and F. Bedard, "Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods," *Agricultural and forest meteorology*, vol. 218, pp. 74- 84, 2016.
- Moosavi AA, Sepaskhah A (2012) Artificial neural networks for predicting unsaturated soil hydraulic characteristics at different applied tensions. *Arch Agron Soil Sci* 58(2):125–153

- N. Kohzadi, B. Milton S, K. Bahman and I. Kaastra, "A comparison of artificial neural network and time series models for forecasting commodity prices," *Neurocomputing*, vol. 10, pp. 169-181, 1996.
- Namin, Sarah Taghavi; Esmaeilzadeh, Mohammad; Najafi, Mohammad; Brown, Tim B and Borevitz, Justin O;, "Deep phenotyping deep learning for temporal phenotypegenotype classification," *Plant methods*, vol. 14, p. 66, 2018.
- Niedbala, G.; Nowakowski, K.; Rudowicz-Nawrocka, J.; Piekutowska, M.; Weres, J.; Tomczak, R.J.; Tyksiński, T.; Pinto, A.Á. Multicriteria prediction and simulation of winter wheat yield using extended qualitative and quantitative data based on artificial neural networks. *Appl. Sci.* 2019, 9, 2773.
- Pandolfi, C.; Mugnai, S.; Azzarello, E.; Bergamasco, S.; Masi, E.; Mancuso, S. Artificial neural networks as a tool for plant identification: A case study on Vietnamese tea accessions. *Euphytica* 2009, 166, 411–421.
- Parmley, K. A., Higgins, R. H., Ganapathysubramanian, B., Sarkar, S., & Singh, A. K. (2019). Machine learning approach for prescriptive plant breeding. *Scientific reports*, 9(1), 17132.
- S. B. Kotsiantis, I. Zaharakis and P. E. Pintelas, "Machine learning: a review of classification and combining techniques," *Artificial Intelligence Review*, vol. 26, pp. 159-190, 2006
- Sadeghi-Tehran, P.; Virlet, N.; Ampe, E.M.; Reyns, P.; Hawkesford, M.J. DeepCount: In-Field Automatic Quantification of Wheat Spikes Using Simple Linear Iterative Clustering and Deep Convolutional Neural Networks. *Front. Plant Sci.* 2019, 10.
- Saini, G.; Khamparia, A.; Luhach, A.K. Classification of Plants Using Convolutional Neural Network. In *Advances in Intelligent Systems and Computing, Proceedings of the First International Conference on Sustainable Technologies for Computational Intelligence*; Luhach, A., Kosa, J., Poonia, R., Gao, X., Singh, D., Eds.; Springer: Singapore, 2020; Volume 1045, pp. 551–561.
- Sastry DVSSR, Upadhyaya HD, Srinivas TR (2019) Variation for seed physical and hydration properties of chickpea (*Cicer arietinum* L) mini core collection and their relevance to conservation and utilization. *Plant Gen Resour* 17:311–324
- Schrider, D.R.; Kern, A.D. Supervised machine learning for population genetics: A new paradigm. *Trends Genet.* 2018, 34, 301–312.
- Shukla, G.K. Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity Edinb.* 1972, 29, 237–245.
- Soares JDR, Pasqual M, Lacerda WS, Silva SO, Donato SLR (2013) Utilization of artificial neural networks in the prediction of the bunches' weight in banana plants. *Sci Hortic* 155:24–29
- Tripodi, P.; Massa, D.; Venezia, A.; Cardi, T. Sensing Technologies for Precision Phenotyping in Vegetable Crops: Current Status and Future Challenges. *Agronomy* 2018, 8, 54.
- Varol, İ. S., Çetin, N., & Kırnak, H. (2022). Evaluation of Image Processing Technique on Quality Properties of Chickpea Seeds (*Cicer arietinum* L.) Using Machine Learning Algorithms. *Journal of Agricultural Sciences*, 29(2), 427-442.
- Wricke, G. Über eine Method ezur Erfassung der ökologischen Streubreitein Feldversuchen. *Z. Pflanzenzuchtg* 1962, 47, 92–96.
- Yan, J., & Wang, X. (2023). Machine learning bridges omics sciences and plant breeding. *Trends in Plant Science*, 28(2), 199-210.
- Yang, H.-W.; Hsu, H.-C.; Yang, C.-K.; Tsai, M.-J.; Kuo, Y.-F. Di erentiating between morphologically similar species in genus *Cinnamomum* (Lauraceae) using deep convolutional neural networks. *Comput. Electron. Agric.* 2019, 162, 739–748.
- Yoosefzadeh Najafabadi, M., Hesami, M., & Eskandari, M. (2023). Machine learning-assisted approaches in modernized plant breeding programs. *Genes*, 14(4), 777.
- Zhao, Jiangan; Bodner, Gernot and Rewald, Boris;, "Phenotyping: using machine learning for improved pairwise genotype classification based on root traits," *Frontiers in plant science*, vol. 7, p. 1864, 2016.