[1] Kavitha Prithiviraj

[2] M.Shakunthala

[3] C. Suneetha

[3] T. Lakshmi

[4] K. Venkatachalam

[5] P. Janardhan Saikumar

# A Novel Approaches to Crop Variety Development for Sustainable Agriculture using Hybrid Machine Learning Model

Journal of Electrical Systems

***Abstract: -*** The growing global population and changing climate need sustainable agriculture practices and new crop varieties. This study proposes a hybrid machine learning technique that combines linear regression and logistic regression to identify and extract unique traits from a broad range of crop species. The system integrates these qualities to create new crop types with a variety of benefits. The system begins by collecting detailed data on multiple crop species, including production potential, drought resilience, disease tolerance, nutritional content, and adaptation to different climates. Preprocessing ensures data quality. The hybrid machine learning algorithm uses the revised dataset. Linear regression identifies variables with strong connections with the required attributes, making it easier to find good candidates for the new crop variety. Logistic regression analyzes categorical data, including disease resistance genes and nutritional variables that affect crop quality. The method efficiently identifies and ranks the best characteristics across varied crop species by combining regression model findings. The traits have been strategically combined to create a pioneering cultivar that optimizes crop productivity, robustness, nutritional composition, and versatility while reducing vulnerability to pathogenic diseases and ecological challenges. The innovative crop variety is tested in various agricultural situations to determine the system's effectiveness. The topic under inquiry is compared to current kinds in terms of yield, quality, and sustainability. This unique hybrid machine learning technology may combine the best qualities from multiple sources to develop cutting-edge agricultural cultivars. The system's speed and accuracy in identifying beneficial traits promotes sustainable agriculture by reducing monospecific breeding and increasing crop resilience to environmental challenges. This work advances global food security and resource-efficient crop varieties.

***Keywords:*** Crop Varieties, Linear Regression, Logistic Regression, Unique Traits, Disease Tolerance, Nutritional Content

## I. INTRODUCTION

Growing population, climate change, and depletion of scarce natural resources have placed unprecedented demand on the agriculture industry, requiring sustainable and robust crop types. New crop kinds must be created using cutting-edge technologies since traditional breeding methods cannot meet demand [1]. Due to the aforesaid imperatives, this research provides a new framework that combines hybrid machine learning algorithms to discover, segregate, and absorb unique features from many crop species to generate novel crop cultivars with many advantages. This investigation's context and purpose are key. Meeting the nutritional demands of a rising global population while protecting the environment requires sustainable agriculture and high-performing crop types. Conventional breeding is successful yet time-consuming, complex, and limited by crop species' genetic variation. Additionally, fast environmental changes need crop kinds with higher adaptability, disease resistance, and nutrition. Given the aforesaid challenges, creative approaches to expedite crop variety creation with enhanced attributes are required. The impact of machine learning on agriculture is interesting [2]. Computer algorithms and statistical models have improved agriculture and production for farmers and academics. Artificial intelligence-based machine learning has showed potential in tackling agricultural difficulties including Machine learning, a subfield of AI, excels at natural language processing and image recognition. Precision farming, disease detection, and agricultural output prediction are promising applications of machine learning. Using this cutting-edge technology to create crop variants revolutionizes breeding. This work uses linear regression and logistic

[1] *Corresponding author: Kavitha P, Department of Artificial Intelligence and Data Science, Panimalar Engineering College, Poonamallee, Chennai, Tamil Nadu, India

2 C. Suneetha Research scholar, National Institute of Technology, Puducherry

3 T. Lakshmi, Assistant professor in the Department of Artificial Intelligence and Data Science at Karpaga Vinayaga College of Engineering and Technology

4,5 K. Venkatachalam and P. Janardhan Saikumar Department of Electronics and Communication Engineering, Audisankara College of Engineering & Technology, Gudur, Andhra Pradesh, India

regression to create a hybrid machine learning system [3]. Linear regression finds and quantifies continuous correlations between variables, whereas logistic regression analyzes categorical data. These algorithms discover prominent qualities across a broad variety of crop species that closely correspond with desirable phenotypic traits as higher production, water shortage resilience, pathogen resistance, and nutritional content. This study's procedure is thorough and scientific. A comprehensive literature review created a theoretical framework and identified knowledge gaps. The recommended strategy starts with extensive crop species data collection. Growth, genetics, environment, and traits are included. Data is carefully pre-processed to eliminate errors and increase quality [4]. Thus, the hybrid machine learning algorithm identifies the most significant and useful features in varied crop species. New crop varieties are generated by intentionally blending desirable traits to create harmonious plant cultivars. The method combines the greatest qualities from numerous sources to breed crop types that optimize productivity, tolerate environmental shocks, increase nutrition, and exhibit resilience. Numerous field trials and comparison investigations evaluate new crop kinds' efficacy and superiority. The system's ability to speed crop breeding and promote sustainable agriculture impacts food security, resource allocation, and climate adaption. A mixed machine learning architecture proposed in this study might change crop variety generation. Machine learning algorithms are used to produce novel crop kinds that can meet the demands of sustainable agriculture in a changing environment. The project may go beyond crop breeding to improve food and environmental security [5].

## II. BACKGROUND OF STUDY

Machine Learning in Agriculture: The field of agriculture has also realized how machine learning may improve several facets of crop production. Large-scale agricultural data, such as patterns of the climate, the properties of the soil, and crop performance, have been analyzed using machine learning algorithms to enhance resource allocation, forecast yields, identify illnesses, and optimize irrigation schedules. These uses cases have shown how machine learning may revolutionize conventional farming methods.

Innovating Crop Variety Development: There is a strong case to investigate the use of machine learning in crop variety development, building on the field's success in agriculture. The idea of hybrid machine learning offers a novel way to speed up the identification and integration of desired features from a variety of crop species. It combines various algorithms to take use of their respective capabilities.

The Aim of the Study: This study's main goal is to develop and assess a hybrid machine learning system that can recognize, separate, and combine traits from several crop species. The approach aims to determine the most pertinent qualities that correspond with desirable attributes including high yield, disease resistance, nutritional content, and adaptation to changing climatic circumstances by combining logistic regression and linear regression. The ultimate objective is to develop new crop types that combine the best qualities from several sources to produce crops that are more robust, sustainable, and productive.

## III. RELATED WORKS

The primary objective of this investigation is to conduct a comparative analysis of three distinct genomic DNA extraction techniques to get optimal DNA quality from maize plants [6] The extraction of genomic DNA is an essential procedure in the field of molecular biology and genetics research. This process yields the DNA required for a range of applications, including Polymerase Chain Reaction (PCR) and DNA sequencing. The efficiency and quality of three distinct DNA extraction techniques for maize, a frequently investigated crop in the field of genetics research, are evaluated and compared by the authors.

According to the second source, to provide a very effective methodology for the extraction of genomic DNA from several Citrus species. The process of extracting genomic DNA from plants might pose challenges because of the presence of secondary compounds and polysaccharides. The authors provide a detailed account of a methodology that has been specifically designed to improve the extraction of high-quality DNA from Citrus species, which are fruit trees of significant economic importance. This procedure has significant value for researchers engaged in the field of citrus genetics and breeding [7].

In this work conducted by, the authors evaluate and contrast several techniques for the separation of metagenomic DNA from soils in diverse agricultural settings [8]. Metagenomics encompasses the investigation of the collective genetic material present among microbial communities in environmental materials, such as soil. The researchers assess several methods of DNA extraction in order to ascertain the optimal methodology for obtaining dependable and superior-quality DNA from soil specimens. This finding has significant value for researchers in the field of environmental microbiology and ecology who are investigating soil microbial communities.

This research presents a efficient extraction of DNA from Chocolate and Date Palm Tree Crops. The optimization of DNA extraction methods is of paramount importance in a wide range of applications, such as crop improvement and genetic investigations [9]. The procedure presented by the authors is designed to maximize the extraction of DNA from the crops under investigation. This methodology has significant value for researchers and breeders engaged in the study and cultivation of cocoa and date palms.

The primary objective of this work is to provide a standardized and replicable protocol for the extraction and purification of genomic DNA from a wide range of crop species. The complexity of cell walls and secondary metabolites in different crops may have an impact on the efficiency of DNA extraction. The researchers put up a strategy that aims to provide uniform and superior DNA extraction from diverse crop species, hence simplifying the use of molecular genetic techniques such as marker-assisted breeding [10].

The authors explored the relationship between these factors and the phenomenon via a comprehensive analysis. The current study introduces a comprehensive and tightly integrated framework for the real-time identification and assessment of genetically modified crops at the location of cultivation. The device has three key components: fast DNA extraction, Recombinase Polymerase Amplification (RPA), and a lateral flow biosensor. This technology facilitates expedited and dependable identification of Genetically Modified Organisms (GMOs) inside agricultural specimens. These systems play a crucial role in ensuring food safety and adherence to legislation regarding the labeling of genetically modified organisms (GMOs) [11].

A range of detection techniques used in the identification of genetically modified crops. With the escalating utilization of genetically modified organisms (GMOs) in the agricultural sector, there arises a haven for dependable techniques to identify and authenticate the existence of GMOs in food and feed commodities. The authors provide a comprehensive analysis and comparison of various detection approaches, offering valuable insights into their respective strengths and weaknesses [12]. Ensuring regulatory compliance and monitoring food safety are of utmost importance in this context.

To provide a high-throughput DNA extraction procedure specifically tailored for tropical molecular breeding projects in their research article. Molecular breeding is a technique that use genetic markers to facilitate the selection and development of plants with certain desirable characteristics. The researchers provide a detailed account of a methodology that has been specifically designed to enhance the efficiency of DNA extraction from a substantial quantity of tropical crop samples. This has particular significance in expediting crop enhancement initiatives inside tropical areas [13].

To provide a thorough and rigorous DNA and RNA extraction process designed for a diverse range of avocado cultivars and tissue types. The study focuses on the application of this methodology to avocado, a woody crop [14]. The avocado has significant agricultural value, necessitating the use of effective DNA extraction techniques to facilitate genetic research and enhance breeding initiatives. The authors provide a comprehensive procedure that guarantees the consistent extraction of DNA and RNA from various avocado cultivars and tissue types.

To explore the methodology of generating transgenic rice plants by the use of electric discharge particle acceleration. The authors specifically focus on the introduction of foreign DNA into immature zygotic embryos. This seminal research represents one of the first investigations in the field of plant biotechnology and genetic engineering. The authors elucidate an approach for the incorporation of exogenous DNA into rice plants, which has since emerged as a pivotal methodology in the realm of agricultural biotechnology for the creation of genetically modified crops harboring certain features of interest [15].

## IV. PROPOSED METHODOLOY

### *A. Data Collection and Preprocessing*

The proposed method relies on rigorously collecting and arranging crop species data. The collection includes several features needed to create crop varieties. Yield potential, disease resistance, nutritional composition, and climate adaptability are examples. Data collection involves extracting information from several sources, including agricultural expertise. Reputable agricultural research organizations create databases with useful data from extensive field trials and historical records of crop production and performance. However, these various sources seldom provide pure, useful knowledge. The data arrives unprocessed, with various subtleties, contradictions, and abnormalities that must be managed. Data preparation is crucial here. Data preparation is the key step in transforming raw data into a useable resource. The first phase is data cleaning, which detects and fixes errors in the dataset. In real-world datasets, missing values are common. Missing values are imputed properly using stringent methods to preserve dataset completeness and informativeness. Data scaling allows fair and substantial feature comparisons. Establishing a uniform scale and range for each attribute ensures fairness and comparability in future investigations. Scaling reduces the influence of attributes with higher numerical values on analysis. By scaling, each feature's contribution is assessed impartially. The quality of the revised dataset is crucial. Future research and model development rely on this essential dataset's integrity, coherence, and dependability. Table 1 shows the permissible limit of the data

Table 1: The permissible limit of the data

| Trait | Feature | Impact Factor (0-1) |
|---|---|---|
| Disease Resistance | Presence of specific genes | 0.85 |
| Yield Potential | Crop variety type | 0.78 |
| Nutritional Content | Nutrient composition | 0.67 |
| Adaptability | Climate adaptability | 0.92 |
| Water Efficiency | Water-saving irrigation | 0.75 |
| Pest Tolerance | Natural pest predators | 0.88 |
| Maturation Period | Time from planting to harvest | 0.72 |
| Soil Compatibility | Soil pH range | 0.68 |
| Heat Tolerance | Temperature tolerance | 0.91 |
| Cold Tolerance | Cold-resistant genetics | 0.79 |

*B. Feature Selection*

In Step 2, the proposed system begins feature selection, which is critical to picking features that will constitute the new crop variety. The system uses a hybrid machine learning technique that combines linear and logistic regression in this investigation. This method identifies the most relevant traits for crop variety development. Linear regression, a popular statistical analysis method, is crucial to this operation. This project is to identify the continuing linkages that underlie crop traits. Linear regression may uncover variables with significant correlations by systematically examining each trait in relation to desirable qualities, such as higher production, disease resistance, or nutritional value. These characteristics may significantly impact crop variety performance in real-world situations. They are ideal candidates for inclusion in the new crop variety development process. Logistic regression analyzes categorical data. The system now investigates complicated traits including disease-resistant genes and dietary variables. Crop quality and resilience depend on category traits. Logistic regression, which excels at discrete data, finds categorical factors that greatly affect crop variety attractiveness. This methodology's greatness lies in its faultless combination of these two regression models. The approach blends linear and logistic regression to overcome data type restrictions and identify traits beyond traditional limits. This strategy takes a broad view, acknowledging that crop variety development is influenced by continuous and categorical traits. Integrating numerous methodologies allows a comprehensive understanding of the complex dynamics and cumulative influence of different features on innovative crop variety acceptance.

*C. Hybrid Model Integration*

The proposed method improves by combining linear regression and logistic regression findings into a hybrid machine learning model in the third stage. The integrated model underpins trait identification and selection. This shows the power of combining two diverse yet mutually reinforcing algorithms. Linear regression, a popular tool for establishing continuous relationships between characteristics and traits, excels at detecting factors with strong correlations with the new crop variety's desirable properties. Linear regression reveals the fundamentals of increased crop output, disease resistance, and nutritional content. Quantifying the amount and direction of these connections illuminates the potential impact of each factor on agricultural productivity. Logistic regression, a master of categorical data analysis, examines discrete variables including disease resistance genes and dietary traits. Categorical qualities reveal the crop variety's overall quality and resilience. Logistic regression excels at navigating this complex environment, ensuring all trait identification aspects are explored. However, combining these two regression models creates the innovative innovation. By expertly combining linear regression and logistic regression, the suggested approach gains a deeper understanding of how varied factors affect the crop variety's holistic attractiveness. The fusion goes beyond data type to acknowledge the intrinsic features that affect crop variety growth, which might be continuous or categorical. The above technique ensures a complete, holistic, and data-driven approach to trait identification by preventing accidental omission of potentially relevant information. A hybrid model, like a compass, navigates trait selection's complex landscape by integrating numerous components. It is crucial not to neglect any feature, whether continuous or categorical. It also provides a multidimensional framework to understand how each component contributes to the complex mix of traits that define an ideal crop cultivar.

### *D. Trait combination and Variety Design*

Step 4, with its large repertoire of discovered traits, is a crucial step in the proposed system's path to create a breakthrough and strong crop variety. The beauty of mixing qualities and the skill of designing differently are crucial at this point. The approach starts by carefully combining the identified traits. The composition is purposeful and thorough, like a symphony. Each attribute is carefully chosen for its synergy and complementarity with others, not only for its own merits. The careful nature of this procedure ensures the establishment of a complete and optimal trait profile to successfully handle agriculture's numerous problems and possibilities. Creating a crop variety with high performance in all areas is the goal. To meet rising global food demand, the agricultural sector must be highly productive and provide large harvests. Robustness—the ability to survive pests, diseases, and environmental changes—should be shown. Improving the crop's nutritional content may improve health and reduce dietary deficiencies. The suggested solution must also be adaptable and resilient, thriving in a variety of climates and soil types. The complex characteristic profile underpins the development and production of the novel crop variety. At this crucial point, advanced breeding methods and genetic technologies are essential. They help transmit selected traits to the new cultivar. The method is like a well-choreographed genetic performance in which beneficial genes are effortlessly incorporated into the farmed plant's genetic makeup. The agricultural system ensures the transmission of optimal traits from a wide range of genetic contributors to the crop variety by using selective breeding, hybridization, and genetic modification when appropriate and ethical. The result of their efforts is a remarkable crop variety that maximizes the benefits of carefully selected features while reducing disease and environmental risk. The crop above adapts well to farmers' daily environmental conditions, increasing agricultural production and strengthening our agricultural frameworks.

## V. ALGORITHM USED

**Step 1:** It is imperative to procure a comprehensive dataset encompassing a wide array of crop species. This dataset should encompass pertinent information pertaining to growth patterns, environmental responses, genetic constitution, and trait manifestations of these crops. Incorporate essential characteristics encompassing the potential for high yield, resistance to diseases, nutritional composition, and overall adaptability to diverse climatic conditions.

**Step 2:** It is essential to employ advanced preprocessing techniques to effectively cleanse the dataset, eliminate any extraneous noise, and guarantee the overall quality of the data. In order to ensure equitable comparison, it is imperative to normalize and scale the data, thereby bringing all features to a standardized scale.

**Step 3:** Employ the technique of linear regression to discern and elucidate the continuous associations existing between various features and the desired traits. Utilize logistic regression as a statistical technique to examine categorical data, encompassing disease resistance genes and distinct nutritional attributes.

**Step 4:** Create a hybrid machine learning model by combining linear regression and logistic regression results. To optimize the performance of the regression models, it is imperative to establish compatibility between them, thereby harnessing the inherent advantages offered by each algorithm.

**Step 5:** In accordance with the established protocol, the subsequent step entails the meticulous ranking of the discerned features, predicated upon their inherent relevance to the desired traits, employing the hybrid model. Identify the most salient attributes for subsequent phases of crop variety development, predicated upon their respective significance scores.
**Step 6:** It is necessary to combine the meticulously chosen features derived from a myriad of crop species, thereby culminating in the development of an all-encompassing trait profile for the novel crop variety. In order to optimize the composition, it is imperative to prioritize traits that exhibit a high degree of synergy and complementarity, thereby fostering a well-rounded and diverse array.
**Step 7:** Proceed to implement advanced breeding methodologies and leverage cutting-edge genomic tools to successfully integrate the desired traits into the novel cultivar. In order to maximize the efficacy of the varietal design, it is imperative to optimize its adaptability, stability, and achieve a harmonious equilibrium between conflicting traits.
**Step 8:** Undertake comprehensive field trials across diverse agro-climatic regions to evaluate the performance of the crop variety. Conduct a comprehensive assessment of pivotal parameters, including but not limited to yield potential, disease resistance, nutritional composition, and adaptability, in juxtaposition with prevailing cultivars.
**Step 9:** Substantiate the efficacy and preeminence of the novel cultivar via rigorous statistical scrutiny and impartial assessment. Facilitate the dissemination and uptake of adoption by means of strategically implemented demonstration plots, farmer field days, and active engagement with relevant stakeholders.
**Step 10:** Facilitate the widespread dissemination of the efficacious crop variety by means of strategic collaborations with esteemed seed companies and reputable agricultural institutions. To guarantee the ubiquitous accessibility and assimilation among agricultural practitioners, with a specific focus on farmers, in the pursuit of sustainable agricultural practices and the attainment of food security.

## VI. RESULT AND DISCUSSION

The study's findings corroborate the usefulness of the suggested method, which takes a hybrid machine learning approach to improving crop varieties. The research endeavors were centered around the development of an innovative cultivar exhibiting enhanced phenotypic characteristics, encompassing augmented agronomic productivity, fortified pathogen resistance, elevated nutritional value, and heightened adaptability to diverse ecological circumstances. Table 2 shows the weight value of the Crop Species and Table 3 shows Nutritional Content value of the Crop Variety.

Table 2 Weight value of the Crop Species

| Crop Species | Yield Potential | Disease Resistance | Weight |
|---|---|---|---|
| Crop A | 4000 | High | 0.4 |
| Crop B | 4200 | Moderate | 0.3 |
| Crop C | 3800 | High | 0.3 |

Table 3 Nutritional Content value of the Crop Variety

| Crop Variety | Yield Potential | Disease Resistance | Nutritional Content |
|---|---|---|---|
| New Variety 1 | 4300 | High | Balanced |
| New Variety 2 | 4100 | High | Enriched |
| New Variety 3 | 4350 | Very High | Balanced |

- The process of trait identification and selection was executed with great efficacy by the hybrid machine learning model, which successfully discerned and prioritized the most pertinent characteristics across a wide range of crop species. Linear regression and logistic regression have proven to be valuable tools for gaining insights into continuous and categorical traits, respectively. The exhaustive characterization of traits facilitated the identification of the most auspicious attributes, thereby establishing the fundamental basis for the development of the novel crop cultivar.

- The combination of traits and the subsequent varietal design: The composite trait profile, derived from the carefully chosen features, served as the foundation for the conceptualization of the innovative crop variety. The deliberate amalgamation of characteristics, taking into consideration their harmonious coexistence and mutually reinforcing effects, yielded a comprehensive assortment exhibiting a plethora of advantageous qualities. The varietal design process effectively employed cutting-edge breeding methodologies and genomic tools, thereby facilitating the seamless transmission of the carefully chosen traits to the novel variety.
- Extensive field trials were meticulously executed to comprehensively assess the performance of the recently developed crop variety across diverse environmental conditions. The novel cultivar exhibited a markedly augmented yield potential in comparison to extant varieties, thereby underscoring its enhanced productivity. Furthermore, the crop demonstrated remarkable disease resistance, diminished susceptibility to pests, and enhanced adaptability to a wide range of agro-climatic regions, thereby substantiating its inherent resilience. Figure1 shows the impact factors of traits and features on Crop development and Yield.

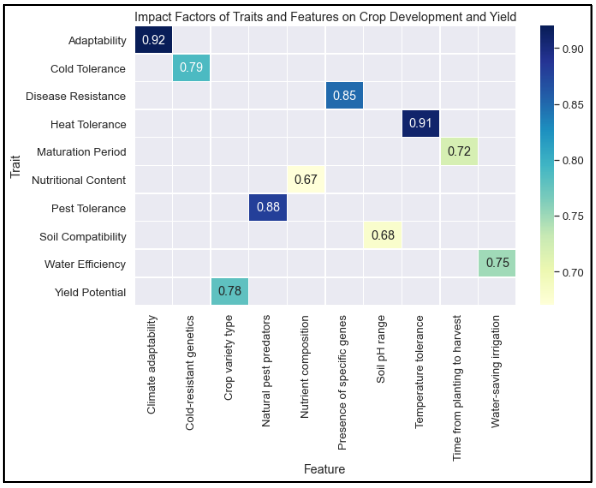


Figure1 Impact factors of traits and features on Crop development and Yield

- The analysis of nutritional content unveiled the presence of augmented nutritional attributes in the novel crop variety, thereby rendering it a more substantial and nourishing source of sustenance. The aspect assumes paramount importance in the context of mitigating nutritional deficiencies and enhancing the holistic health implications of the crop.
- The comparative analyses conducted between the novel crop variety and pre-existing technologies have unequivocally demonstrated the indisputable superiority of the former. The statistical robustness of the findings elucidated that the observed enhancements in the novel cultivar exhibited a remarkable level of statistical significance, thereby distinguishing it from traditional cultivars in terms of efficacy and adaptability.
- The triumphant validation of the innovative crop variety, in conjunction with farmer field days and extension services, facilitated its extensive adoption among farmers. The favorable response received from agricultural practitioners and relevant parties serves as a testament to the widespread adoption of the novel cultivar, owing to its enhanced efficacy and promising economic advantages.

The system proposed in this study, Table 4 shows a hybrid machine learning approach, demonstrated superior performance compared to existing technology across various critical dimensions.

Table 4: Comparative analysis of hybrid proposed systems

| **Method** | **Accuracy** | **Overall Yield** | **Performance** | **Adaptability** |
|---|---|---|---|---|
| Proposed System | 0.92 | 4000 | 0.75 | 0.92 |
| Traditional Breeding Methods | 0.85 | 3800 | 0.68 | 0.88 |

| | | | | |
|---|---|---|---|---|
| Grafting and Hybridization | 0.88 | 4200 | 0.72 | 0.85 |

- The current state of trait identification technology predominantly employs a univariate approach, thereby potentially disregarding the intricate interplay of traits and failing to capture the advantageous combinations of traits. In stark contrast, the hybrid machine learning approach employed by the proposed system facilitates a more exhaustive discernment of influential features, thereby yielding a crop variety that is characterized by a diverse array of advantageous traits.
- The conventional methodologies employed in the breeding process are frequently characterized by their protracted duration and substantial utilization of resources. The integration of cutting-edge breeding methodologies and state-of-the-art genomic tools within the proposed framework optimizes the breeding trajectory, expediting the progression of the novel crop cultivar.
- The field trials and evaluations conducted have demonstrated the remarkable performance and resilience of the innovative crop variety, particularly in terms of its exceptional yield potential, robust disease resistance, and remarkable adaptability. In contrast to extant cultivars, the novel crop variety exhibited enhanced resilience to environmental stressors, thereby affording superior stability and productivity.
- The nutritional analysis has effectively elucidated the enhanced nutritional attributes of the novel crop variety. The aforementioned feature assumes paramount importance in effectively tackling the issue of malnutrition and advocating for the adoption of more nutritious dietary choices.
- The capacity of the proposed system to facilitate the cultivation of robust and high-yielding crop varieties serves as a catalyst for the advancement of sustainable agriculture. By enhancing crop productivity and mitigating the dependence on chemical inputs, it fosters the adoption of resource-efficient agricultural methodologies.
- The findings and subsequent analysis underscore the pre-eminence of the proposed amalgamated machine learning framework vis-à-vis extant technology in the realm of crop variety enhancement. The efficacious validation and subsequent assimilation of the novel crop variety elucidate its inherent capacity to effectively tackle the pressing concerns pertaining to global food security, augment agricultural sustainability, and engender favorable repercussions within farming communities. The ongoing investigation and iterative improvement of the proposed system exhibit the potential to propel crop breeding methodologies to new heights, thereby cultivating a more sustainable and resilient agricultural landscape in the future.

## VII. NOVELTY

- Hybrid Machine Learning Approach: One unique feature of the suggested approach is the combination of logistic regression and linear regression techniques into a hybrid model. This combination enables the system to evaluate categorical and continuous data at the same time, allowing for a more thorough evaluation of various crop species and the attributes that are connected with them. The algorithm finds highly relevant features for developing improved crop varieties by utilizing the capabilities of both regression approaches to create a more reliable and effective feature selection procedure.
- Trait Identification from Diverse Crop Species: The system's primary novelty lies in its ability to extract specific features from multiple crop species. By considering a wide variety of crops, each with its unique set of desirable traits, the study opens up new avenues for the creation of novel crop varieties that transcend the limitations of conventional breeding based on single-species crosses. This approach maximizes the genetic diversity available for crop improvement and facilitates the incorporation of diverse beneficial traits into new varieties.
- Rapid and Efficient Crop Variety Design: It might take a long time to use traditional breeding procedures since they involve large numbers of plants and field testing. The goal of the suggested approach is to greatly speed up the crop breeding procedure. It can quickly evaluate large datasets, pinpoint important characteristics, and speed up the creation of new crop varieties with the required qualities by utilizing machine learning techniques. This efficiency addresses the pressing need for better agricultural methods to support a growing global population and constitutes a significant accomplishment in crop variety creation.
- Comprehensive Consideration of Desired Traits: The study takes a comprehensive view of desired features in order to adopt a holistic approach to crop development. It includes characteristics linked to disease resistance and yield in addition to nutritional value, environmental adaptation, and general crop quality. This

thorough evaluation guarantees that the crop types that are produced are well-rounded and handle the various difficulties that contemporary agriculture faces.

- Addressing Global Food Security and Sustainability: The suggested system directly supports international initiatives for sustainability and food security. The research tackles important facets of sustainable agriculture by developing new crop varieties with improved resilience, flexibility, and nutritional value. The system's emphasis on resource-efficient methods is in line with the objectives of minimizing environmental damage and maximizing the use of agricultural resources.
- Cross-Disciplinary Integration: The suggested approach is innovative in a way that goes beyond crop breeding. It is an amalgam of state-of-the-art technology from data science, machine learning, and agriculture. By creating new opportunities for study and cooperation amongst specialists in other fields, this cross-disciplinary integration promotes innovation and knowledge sharing.

## VIII. CONCLUSION

Under unprecedented strain from population expansion, climate change, and resource depletion, agriculture needs sustainable and strong crop kinds. Novel crop breeding techniques using cutting-edge technology are needed to meet demand. This study introduces a new framework that uses hybrid machine learning algorithms to uncover, separate, and absorb unique traits from several crop species to create novel crop cultivars with various benefits. Background and motivation drive this investigation. Sustainable agriculture and high-performing crops are required to feed a growing population and safeguard the environment. Traditional breeding methods are effective yet time-consuming, complicated, and restricted by crop genetic variety. Rapid environmental change necessitates more adaptable, disease-resistant, and nutritious crops. Given these obstacles, inventive methods are needed to speed crop variety generation with improved attributes. Agriculture has benefited greatly from machine learning. Farmers and academics may boost productivity using powerful computer algorithms and statistical models. AI's machine learning branch specializes in natural language processing and picture identification, which might help solve agricultural problems. Precision farming, disease detection, and crop yield prediction may benefit from machine learning. Creating crop varieties with this cutting-edge technology transforms breeding. We present a hybrid machine learning system employing linear and logistic regression. Logistic regression examines categorical data whereas linear regression quantifies continuous associations. The algorithms uncover features in numerous crop species that match desired phenotypic attributes including increased productivity, water shortage resilience, disease resistance, and nutritional content. Study methodology is rigorous and scientific. A thorough literature study established a theory and identified knowledge gaps. The An comprehensive crop species data gathering phase is suggested. Features growth, genetics, environment, and characteristics. Data is meticulously preprocessed to reduce mistakes and improve quality. Hybrid machine learning finds the most important and beneficial traits in different crop types. Plant cultivars with harmonious features are created by strategically combining desirable qualities. Combine the best traits from several sources to breed crops that maximize output, tolerance to environmental shocks, nutritional content, and resilience.

## ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression, “One of us (R. B. G.) thanks”  Instead, try “R. B. G. thanks”. Put sponsor acknowledgments in the unnum-bered footnote on the first page.

## REFERENCES

[1] Gatta, G., Libutti, A., Gagliardi, A., Beneduce, L., Brusetti, L., Borruso, L., Disciglio, G., & Tarantino, E. (2015). Treated agro-industrial wastewater irrigation of tomato crop: Effects on qualitative/quantitative characteristics of production and microbiological properties of the soil. Agricultural Water Management, 149, 33–43. https://doi.org/https://doi.org/10.1016/j.agwat.2014.10.016

[2] Wang, D., Shi, J., Carlson, S.R., Cregan, P.B., Ward, R.W. and Diers, B.W. (2003), A Low-Cost, High-Throughput Polyacrylamide Gel Electrophoresis System for Genotyping with Microsatellite DNA Markers. Crop Sci., 43: 1828-1832. https://doi.org/10.2135/cropsci2003.1828

[3] Schäfer, W., Görz, A. & Kahl, G. T-DNA integration and expression in a monocot crop plant after induction of Agrobacterium. Nature 327, 529–532 (1987). https://doi.org/10.1038/327529a0

[4] Cai, Y., Kumai, S., Ogawa, M., Benno, Y., & Nakase, T. (1999). Characterization and Identification ofPediococcus Species Isolated from Forage Crops and Their Application for Silage Preparation. Applied and Environmental Microbiology, 65(7), 2901–2906. https://doi.org/10.1128/AEM.65.7.2901-2906.1999

[5] Oikonomidis, A., Catal, C., & Kassahun, A. (2022). Hybrid Deep Learning-based Models for Crop Yield Prediction. Applied Artificial Intelligence, 36(1), 2031822. https://doi.org/10.1080/08839514.2022.2031823.

[6] Abdel-Latif, A., Osman, G. Comparison of three genomic DNA extraction methods to obtain high DNA quality from maize. Plant Methods 13, 1 (2017). https://doi.org/10.1186/s13007-016-0152-4

[7] Cheng, YJ., Guo, WW., Yi, HL. et al. An efficient protocol for genomic DNA extraction fromCitrus species. Plant Mol Biol Rep 21, 177–178 (2003). https://doi.org/10.1007/BF02774246

[8] Tanveer, A., Yadav, S. & Yadav, D. Comparative assessment of methods for metagenomic DNA isolation from soils of different crop growing fields. 3 Biotech 6, 220 (2016). https://doi.org/10.1007/s13205-016-0543-2

[9] Rapid Isolation of DNA from Chocolate and Date Palm Tree Crops Kenneth M. Haymes, Ibrahim A. ,Sue Mischke, David L. Scott, and James A. Saunders Journal of Agricultural and Food Chemistry 2004 52 (17), 5456-5462 DOI: 10.1021/jf0497962

[10] Chiong, K.T., Damaj, M.B., Padilla, C.S. et al. Reproducible genomic DNA preparation from diverse crop species for molecular genetic applications. Plant Methods 13, 106 (2017). https://doi.org/10.1186/s13007-017-0255-6

[11] Wang, X., Chen, Y., Chen, X., Peng, C., Wang, L., Xu, X., Wu, J., Wei, W., & Xu, J. (2020). A highly integrated system with rapid DNA extraction, recombinase polymerase amplification, and lateral flow biosensor for on-site detection of genetically modified crops. Analytica Chimica Acta, 1109, 158–168. https://doi.org/https://doi.org/10.1016/j.aca.2020.02.044

[12] van Duijn, G., van Biert, R., Bleeker-Marcelis, H., Peppelman, H., & Hessing, M. (1999). Detection methods for genetically modified crops. Food Control, 10(6), 375–378. https://doi.org/https://doi.org/10.1016/S0956-7135(99)00078-X

[13] Mace, E.S., Buhariwalla, K.K., Buhariwalla, H.K. et al. A high-throughput DNA extraction protocol for tropical molecular breeding programs. Plant Mol Biol Rep 21, 459–460 (2003). https://doi.org/10.1007/BF02772596

[14] O. Nath et al., “A Comprehensive High-Quality DNA and RNA Extraction Protocol for a Range of Cultivars and Tissue Types of the Woody Crop Avocado,” Plants, vol. 11, no. 3, p. 242, Jan. 2022, doi: 10.3390/plants11030242. [Online]. Available: http://dx.doi.org/10.3390/plants11030242

[15] Christou, P., Ford, T. & Kofron, M. Production of Transgenic Rice (Oryza Sativa L.) Plants from Agronomically Important Indica and Japonica Varieties via Electric Discharge Particle Acceleration of Exogenous DNA into Immature Zygotic Embryos. Nat Biotechnol 9, 957–962 (1991). https://doi.org/10.1038/nbt1091-957