Agri Articles
(e-Magazine for Agricultural Articles)
Volume: 06, Issue: 04 (JULY-AUG, 2026)
Available online at http://www.agriarticles.com

AI Revolution in Plant Breeding: From Field to Future

*Saurav Raj[1], Madhur Kokane[2], Rohit Malviya[3] and Kanshika Vaibhav[4]
[1]Department of Genetics and Plant Breeding, Rajmata Vijayaraje Scindia Krishi Vishwa Vidyalaya, R.A.K College of Agriculture, Sehore, Madhya Pradesh India
[2]Department of Genetics and Plant Breeding, Lakshmi Narain College of Technology and Science, Bhopal, Madhya Pradesh India
[3]Ph.D. Research Scholar, Department of Genetics and Plant Breeding, Rajmata Vijayaraje Scindia Krishi Vishwa Vidyalaya, Gwalior, Madhya Pradesh, India
[4]Department of Molecular Biology and Biotechnology, Rani Lakshmi Bai Central Agricultural University Jhansi, Uttar Pradesh, India
*Corresponding Author's email: sauravraj766@gmail.com

Every extra quintal of grain harvested from a hectare of Indian soil today has to work harder than it did a generation ago. The population that depends on that harvest keeps climbing, the monsoon keeps behaving more erratically, and the land available for expansion keeps shrinking. Plant breeders have always answered such pressure with patience: crossing, selecting, testing across seasons, and waiting for genetics to reveal itself in the field. That patience is a luxury the world can no longer fully afford.

Climate change has added volatility that older breeding pipelines were never designed to absorb. Heat waves during flowering, erratic rainfall, and newly emerging pest and pathogen pressures are shifting faster than a conventional ten-year varietal development cycle can track. Climate variability, economic shocks and conflict are compounding to threaten the sustainable production of nutritious food, and researchers have flagged that hunger and malnutrition remain stubbornly high despite decades of agricultural progress (Khan et al., 2023). Meeting the 2030 goals on food security, in other words, cannot rely on doing what breeders have always done, only slightly faster.

This is precisely the gap into which artificial intelligence has stepped. Not as a replacement for the breeder's eye and judgement, but as an amplifier of it sorting through genomic markers, drone images and weather layers at a scale no human team could manage unaided, and returning predictions a breeder can act on in the same season rather than the next one. What was once a slow accumulation of field notebooks is becoming a live feed of trait predictions, disease alerts and genomic breeding values. This article walks through what that shift actually looks like on the ground, from the statistics behind genomic selection to the drones flying over experimental plots, and asks, honestly, where the technology still falls short.

Artificial Intelligence in Plant Breeding

Artificial intelligence, at its simplest, is a computing system that improves its own performance on a task by learning from data rather than following a fixed set of rules written in advance. Machine learning is the part of AI most relevant to breeding: algorithms are shown thousands of examples - say, genetic markers paired with yield outcomes - and they learn the statistical relationship between the two well enough to predict yield in genotypes they have never seen. This is a genuine departure from conventional data analysis, which typically tests a small number of pre-specified relationships (does trait A correlate with trait B) rather than discovering unexpected patterns buried in high-dimensional data.

Deep learning is a further refinement built from layered neural networks that can model relationships too complex or non-linear for classical statistics to capture cleanly. Multilayer perceptrons and convolutional neural networks, for instance, have been used directly on genomic marker data structured as an "image," achieving prediction accuracies competitive with - and in large datasets, better than - the standard GBLUP statistical model used in genomic selection for decades (Crossa et al., 2025). Computer vision, a specialised branch of deep learning, teaches a model to interpret pixels: to recognise a lesion on a leaf, count tillers in a plot, or estimate canopy cover from a photograph, tasks that previously required a trained human scout walking every row.

None of this works without data, and modern breeding programmes now generate an extraordinary volume of it. A single genomic selection trial can involve tens of thousands of SNP markers screened across thousands of lines; a single drone flight over a field trial can produce gigabytes of multispectral imagery in an afternoon. Big data in agriculture refers to this scale and variety of information - genomic, phenomic, environmental and now increasingly textual, drawn from scientific literature itself - and AI is essentially the toolkit built to make sense of it before a single season is lost to indecision.

AI Applications in Modern Plant Breeding

High-Throughput Phenotyping: Phenotyping - measuring what a plant actually looks like and how it performs - has long been the bottleneck in breeding, trailing far behind the speed at which genotyping technology has advanced. High-throughput phenotyping (HTP) closes that gap using drones, satellites, RGB cameras and hyperspectral sensors to capture plant traits non-destructively and at scale. Unmanned aerial vehicles equipped with RGB, multispectral or hyperspectral sensors can now measure plant height, canopy cover, biomass, senescence and even early disease symptoms across an entire trial in a single flight, feeding that data directly into breeding decisions (Gano et al., 2024). Ground-based and aerial HTP platforms together have substantially reduced the phenotyping bottleneck that once limited the pace of genetic gain in breeding programmes (Gill et al., 2022). Figure 1 illustrates how artificial intelligence integrates field phenotyping, genomic data, environmental information, and predictive analytics to accelerate variety development. The real payoff comes when this imagery is combined with genomic data rather than used in isolation. Field trials that layered drone-derived phenomic data on top of genomic information markedly improved prediction accuracy for grain yield, test weight and grain protein content in winter wheat, in some cases by tens of percentage points over genomics alone (Gill et al., 2024). That combination - phenomics feeding genomics - is quietly becoming one of the more consequential ideas in modern breeding.

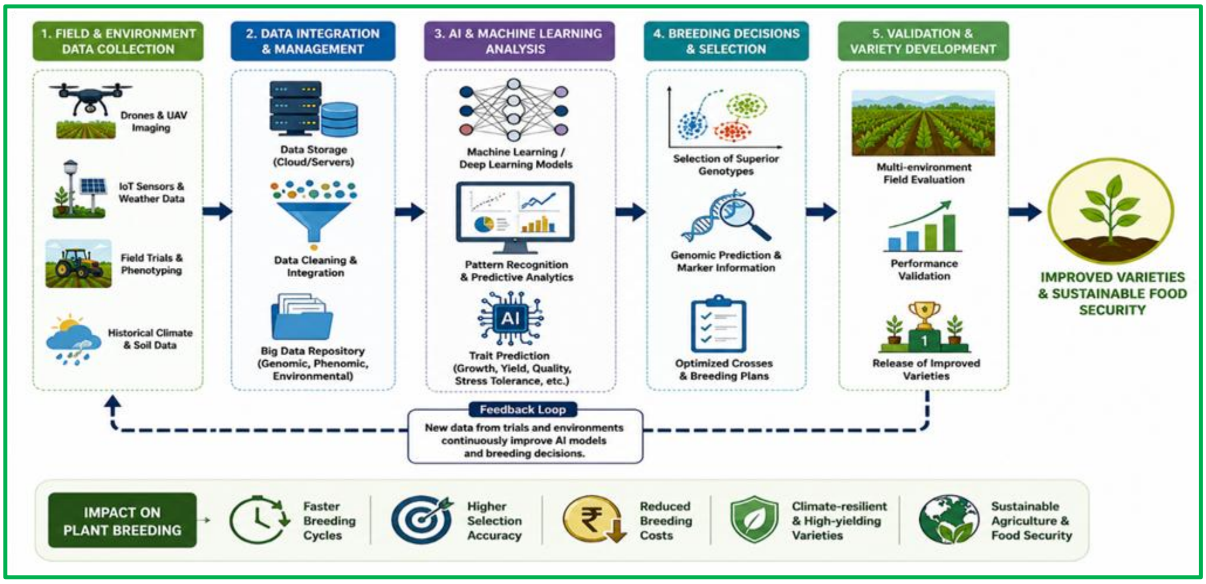

Figure 1. Artificial Intelligence Pipeline in Modern Plant Breeding

Genomic Selection: Genomic selection uses genome-wide marker data to predict the breeding value of a candidate line before it is ever planted in a yield trial, allowing breeders to discard poor performers and advance promising ones years earlier than phenotypic selection alone would permit. Two decades after it was first proposed, genomic selection has become a mainstream tool precisely because machine learning and deep learning methods have made its predictions more accurate and more affordable to generate, especially as training population size, marker density and multi-environment data have all grown (Alemu et al., 2024). Neural-network-based genomic prediction models, tested across wheat, maize and tomato datasets, have matched or exceeded the accuracy of established statistical models while running at comparable or faster computational speed (Crossa et al., 2025). For students preparing for competitive examinations, the practical takeaway is this: genomic selection does not replace field trials, it triages which lines deserve one. That triage is what compresses a breeding cycle from a decade to considerably less.

Disease and Pest Detection: Image recognition trained on leaf photographs can now flag fungal, bacterial and viral symptoms often before visible damage would prompt a farmer or scout to raise an alarm. A systematic review of deep learning approaches to plant disease detection found that convolutional neural network architectures consistently outperformed manual scouting on accuracy and speed, while also lowering the expertise threshold needed to make a diagnosis in the field (Shoaib et al., 2023). This matters as much for extension services reaching smallholders as it does for breeding programmes screening thousands of genotypes for resistance under artificial inoculation.

Precision Agriculture: Precision agriculture applies the same predictive logic to field management rather than genetics: smart irrigation systems that respond to real-time soil moisture and canopy stress, variable-rate nutrient application guided by remote sensing, and yield forecasting models that combine satellite imagery, weather data and soil properties. Machine-learning-based yield prediction has grown rapidly as a research area since 2021, with random forest, gradient boosting and hybrid deep-learning frameworks becoming the standard toolkit for translating remote-sensing and weather data into actionable yield forecasts. These tools increasingly inform not just farm-level decisions but also which environments a breeding programme should prioritise for multi-location trials.

Speed Breeding and AI: Speed breeding shortens the breeding cycle itself by manipulating photoperiod, light quality and temperature to force multiple plant generations through in a single year - as many as six generations annually for spring wheat under controlled conditions, compared with two or three in an ordinary glasshouse (Watson et al., 2018). On its own, speed breeding accelerates generation turnover; paired with AI-driven genomic selection, it compounds that advantage, because breeders can now decide which of those rapidly advancing generations are worth carrying forward using a predicted breeding value rather than waiting for a full phenotypic evaluation. Reviews of the technique note that its main constraints are not biological but infrastructural - the cost of controlled-environment facilities and the trained personnel needed to run them - precisely the kind of constraint that decision-support software and predictive models can help a programme use more efficiently (Wanga et al., 2021).

Benefits of AI in Crop Improvement

Table 1 compares conventional and AI-assisted breeding approaches across the parameters that matter most to a breeding programme manager: time, accuracy, cost, and the underlying quality of decision making.

Table 1. Traditional Plant Breeding vs AI-Assisted Plant Breeding

Parameter	Conventional Breeding	AI-Assisted Breeding
Time to develop a variety	8–12 years, dependent on generation cycles	4–7 years, aided by predictive models and speed breeding
Selection accuracy	Based largely on visual scoring and field trials	Higher, driven by genomic prediction and image-based traits

Parameter	Conventional Breeding	AI-Assisted Breeding
Cost per cycle	High recurring cost for multi-season field trials	Higher upfront investment, lower long-term cost per gain
Data analysis	Manual, spreadsheet-based statistics	Automated, handles large multi-omics datasets
Phenotyping	Manual scoring, slow and labour intensive	Drone, sensor and image-based, rapid and scalable
Selection efficiency	Moderate, limited by human observation	High, supported by predictive breeding values
Climate resilience	Reactive, based on past seasonal experience	Proactive, guided by climate and stress-prediction models
Decision making	Breeder judgement and experience	Data-driven, decision-support assisted

The comparison in Table 1 makes clear that AI-assisted breeding is not simply "faster" in a vague sense; it is faster because each stage of the pipeline - phenotyping, prediction, selection - becomes less dependent on waiting for a full season to pass before a decision can be made. The reduction in breeding cycle length, in turn, is what shortens the path from a promising cross to a variety a farmer can actually plant, which is where climate resilience and resource-use efficiency become tangible outcomes rather than aspirations. Genomic selection studies consistently report that predicting genetic merit early, and combining it with tools that shorten generation time, is what most directly increases the rate of genetic gain a programme can deliver per year (Alemu et al., 2024). Decision support is perhaps the least visible but most pervasive benefit. Breeders juggling hundreds of crosses across multiple locations and years increasingly rely on software that ranks candidates, flags anomalies in phenotypic data, and recommends which environments to prioritise - freeing scientific judgement for the decisions that genuinely require it, such as interpreting an unexpected result or setting breeding objectives in the first place.

Challenges

None of this comes without friction. Machine learning models are only as good as the data used to train them, and many breeding programmes, particularly in the public sector across the developing world, still lack the large, well-curated, multi-year datasets that deep learning methods need to perform reliably. Reviews of genomic prediction repeatedly emphasise that prediction accuracy depends heavily on training population size and its genetic relatedness to the population being selected - a data requirement that is expensive to satisfy and not evenly distributed across crops or countries (Crossa et al., 2025).

Computational cost is a second constraint. Deep learning models, especially those handling genomic, image and environmental data simultaneously, require processing power and storage that many public breeding institutions in low- and middle-income countries do not have on hand, even if the underlying algorithms themselves are freely available. This creates a real risk that AI-assisted breeding could widen rather than narrow the gap between well-resourced multinational programmes and public breeding systems working with tighter budgets.

A related shortage is human, not computational: breeding programmes need scientists who are genuinely fluent in both quantitative genetics and machine learning, and that combination of skills remains rare enough that many institutions default to hiring one or the other rather than training for both. Ethical and governance questions are also emerging as AI models increasingly assist decisions such as which germplasm to prioritise or which markers count as functionally significant - decisions that were previously made through peer-reviewed scientific consensus and are now sometimes made, or at least pre-filtered, by a model whose reasoning is not always transparent to the breeder using it. Finally, adoption at the smallholder level remains constrained by patchy rural infrastructure, uneven digital literacy, and the practical difficulty of getting a drone-derived recommendation to a farmer who may not have reliable internet access to receive it. Table 2 summarises the major AI technologies currently in use across plant breeding, alongside their principal applications and advantages.

Table 2. Major Applications of Artificial Intelligence in Plant Breeding

AI Technology	Plant Breeding Application	Advantages	Example
Machine Learning	Genomic prediction of breeding values	Faster, more accurate selection	Genomic selection in wheat and maize
Deep Learning	Multi-omics data integration	Captures complex, non-linear patterns	Neural network genomic prediction models
Computer Vision	Disease and stress detection from images	Early, non-destructive diagnosis	CNN-based leaf disease classifiers
Genomic Selection	Prediction of complex quantitative traits	Shortens breeding cycle length	Yield and quality prediction in cereals
Remote Sensing	Field-level phenotyping and yield mapping	Scalable, rapid data capture	Drone and satellite crop monitoring
Robotics	Automated planting, sampling and sorting	Reduces labour, increases precision	Autonomous phenotyping platforms

Future Prospects

The integration of AI with genome editing tools such as CRISPR-Cas is arguably where the next major leap in precision will come from. Deep learning models are already used to design guide RNAs, predict on-target efficiency and flag likely off-target effects with a precision that manual sequence inspection cannot match, and reviews of the field expect this integration to deepen as more editing outcome data accumulates (Dixit et al., 2024; Lee, 2023). Coupling AI-guided genome editing with genomic selection and speed breeding - target identification, precise editing and rapid generation advance operating as a single accelerated pipeline - is the clearest example of how these individually useful technologies compound when stacked together.

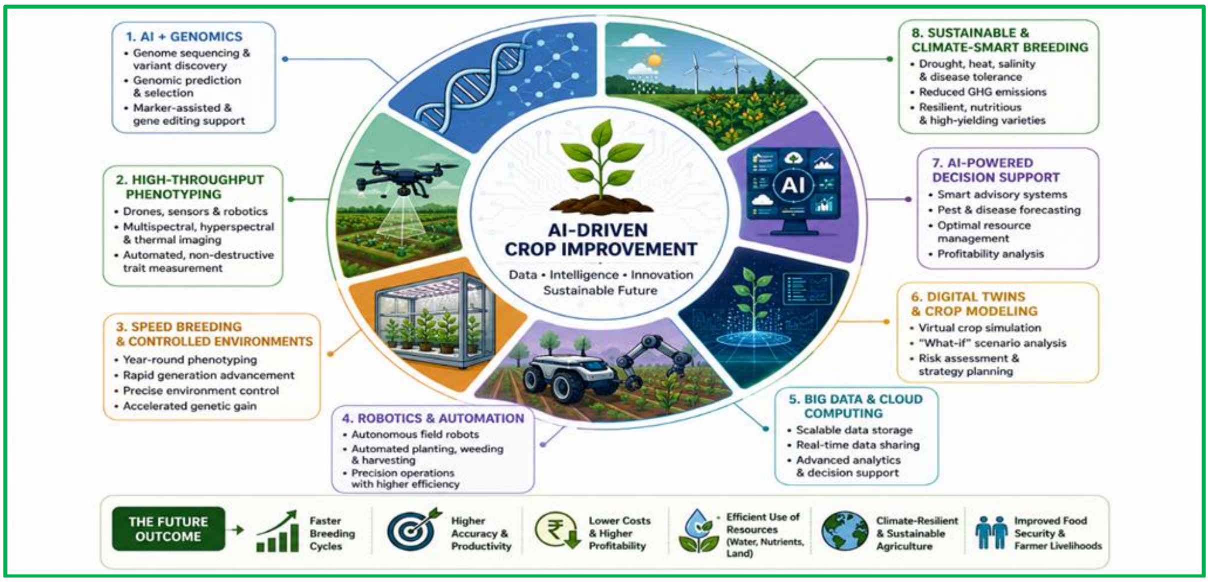

Figure 2. Future Roadmap of AI-Driven Crop Improvement

Digital twins of crops and breeding trials, in which a virtual model of a genotype's performance is simulated under different environmental scenarios before it is ever planted, are moving from a research concept toward practical breeding-support tools. Robotics is following a similar path in phenotyping platforms, gradually automating the physically repetitive parts of field data collection that currently consume disproportionate researcher time. A newer development worth watching closely is the emergence of large language models trained directly on plant genomic sequences, which are beginning to function as general-purpose predictive engines for gene expression, regulatory elements and functional variants across dozens of crop species simultaneously, rather than requiring a bespoke model built from scratch for every trait and species (Lam et al., 2024).

Perhaps the most consequential future application, particularly for a country like India, lies in orphan and underutilised crops - millets, legumes and regional staples that have historically received a fraction of the genomic and breeding investment that wheat, rice and maize have. Machine learning approaches that transfer knowledge learned from well-studied major crops onto these orphan crops are now being explored specifically because they promise to shrink the historical resource gap without requiring decades of dedicated genomic infrastructure to be built from scratch for each neglected species (MacNish et al., 2025). For nutritionally important but long-neglected Indian staples, that transfer-learning approach could matter as much as any single technological breakthrough.

Conclusion

Artificial intelligence has moved from being a novelty in plant breeding conferences to a working part of the breeding pipeline: predicting breeding values before a seed is sown, spotting disease before the human eye would catch it, and compressing a decade of generation advance into a handful of years. None of this displaces the breeder's core scientific judgement - deciding what traits actually matter, interpreting a genuinely surprising result, or setting the direction of a programme still rests with people, not models. What AI changes is how much of the routine, data-heavy work stands between a good idea and a variety in a farmer's field. Used well, and paired honestly with continued field evaluation rather than as a substitute for it, that is a meaningful acceleration for a discipline that food security can no longer afford to leave running at its traditional pace.

As shown in Table 2, no single AI technology covers the entire breeding pipeline; machine learning and genomic selection handle prediction from marker data, computer vision and remote sensing handle phenotyping, and robotics is beginning to automate the physical labour connecting the two. The real strength of an AI-assisted breeding programme lies in stitching these technologies together rather than adopting any one in isolation.

References

1. Alemu, A., Åstrand, J., Montesinos-López, O. A., Isidro y Sánchez, J., Fernández-Gónzalez, J., Tadesse, W., Vetukuri, R. R., Carlsson, A. S., Ceplitis, A., Crossa, J., Ortiz, R., & Chawade, A. (2024). Genomic selection in plant breeding: Key factors shaping two decades of progress. Molecular Plant, 17(4), 552–578. https://doi.org/10.1016/j.molp.2024.03.007
2. Cheng, Q., & Wang, X. (2024). Machine learning for AI breeding in plants. Genomics, Proteomics & Bioinformatics, 22. https://doi.org/10.1093/gpbjnl/qzae051
3. Crossa, J., Martini, J. W. R., Vitale, P., Pérez-Rodríguez, P., Costa-Neto, G., Fritsche-Neto, R., Runcie, D., Cuevas, J., Toledo, F., Li, H., De Vita, P., Gerard, G., Dreisigacker, S., Crespo-Herrera, L., Saint Pierre, C., Bentley, A., Lillemo, M., Ortiz, R., Montesinos-López, O. A., & Montesinos-López, A. (2025). Expanding genomic prediction in plant breeding: Harnessing big data, machine learning, and advanced software. Trends in Plant Science, 30(3), 268–284. https://doi.org/10.1016/j.tplants.2024.12.009
4. Dixit, S., Kumar, A., Srinivasan, K., Vincent, P. M. D. R., & Ramu Krishnan, N. (2024). Advancing genome editing with artificial intelligence: Opportunities, challenges, and future directions. Frontiers in Bioengineering and Biotechnology, 11, 1335901. https://doi.org/10.3389/fbioe.2023.1335901
5. Esposito, S., Ruggieri, V., & Tripodi, P. (2022). Editorial: Machine learning for big data analysis: Applications in plant breeding and genomics. Frontiers in Genetics, 13, 916462. https://doi.org/10.3389/fgene.2022.916462
6. Gano, B., Bhadra, S., Vilbig, J. M., Ahmed, N., Sagan, V., & Shakoor, N. (2024). Drone-based imaging sensors, techniques, and applications in plant phenotyping for crop

breeding: A comprehensive review. The Plant Phenome Journal, 7(1), e20100. https://doi.org/10.1002/ppj2.20100

7. Kaushal, S., Gill, H.S., Billah, M.M., Khan, S.N., Halder, J., Bernardo, A., Amand, P.S., Bai, G., Glover, K., Maimaitijiang, M. and Sehgal, S.K. (2024). Enhancing the potential of phenomic and genomic prediction in winter wheat breeding using high-throughput phenotyping and deep learning. Frontiers in Plant Science, 15, 1410249. https://doi.org/10.3389/fpls.2024.1410249
8. Gill, T., Gill, S. K., Saini, D. K., Chopra, Y., de Koff, J. P., & Sandhu, K. S. (2022). A comprehensive review of high throughput phenotyping and machine learning for plant stress phenotyping. Phenomics, 2(3), 156–183. https://doi.org/10.1007/s43657-022-00048-z
9. Khan, M. I. R., Irfan, M., & Gupta, R. (2023). Editorial: Improving crop nutritional security for sustainable agriculture in the era of climate change. Frontiers in Plant Science, 14, 1292264. https://doi.org/10.3389/fpls.2023.1292264
10. Lam, H. Y. I., Ong, X. E., & Mutwil, M. (2024). Large language models in plant biology. Trends in Plant Science, 29(10), 1145–1155. https://doi.org/10.1016/j.tplants.2024.04.013
11. Lee, M. (2023). Deep learning in CRISPR-Cas systems: A review of recent studies. Frontiers in Bioengineering and Biotechnology, 11, 1226182. https://doi.org/10.3389/fbioe.2023.1226182
12. MacNish, T. R., Danilevicz, M. F., Bayer, P. E., Bestry, M. S., & Edwards, D. (2025). Application of machine learning and genomics for orphan crop improvement. Nature Communications, 16, 982. https://doi.org/10.1038/s41467-025-56330-x
13. Shoaib, M., Shah, B., EI-Sappagh, S., Ali, A., Ullah, A., Alenezi, F., Gechev, T., Hussain, T., & Ali, F. (2023). An advanced deep learning models-based plant disease detection: A review of recent research. Frontiers in Plant Science, 14, 1158933. https://doi.org/10.3389/fpls.2023.1158933
14. Wanga, M. A., Shimelis, H., Mashilo, J., & Laing, M. D. (2021). Opportunities and challenges of speed breeding: A review. Plant Breeding, 140(2), 185–194. https://doi.org/10.1111/pbr.12909
15. Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S. S., James, R., Mendoza-Suárez, M., Ramirez-Gonzalez, R. H., Chinchilla, S. P., ... Hickey, L. T. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. Nature Plants, 4(1), 23–29. https://doi.org/10.1038/s41477-017-0083-8