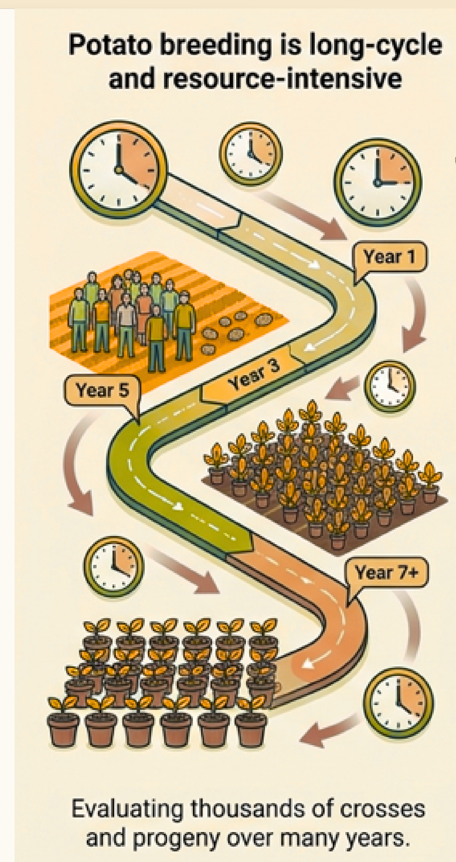
A MULTI-OBJECTIVE NSGA-LL APPROACH TO OPTIMISING GENOMIC AND PEDIGREE-BASED POTATO BREEDING

Ollscoil Teicneolaíochta an Oirdheiscirt
South East Technological University

Daksh Yogesh Prajapati | Supervisor: Dr Rejwanul Haque | MSc in Applied AI, SETU Carlow

INTRODUCTION

Potato breeding: long-cycle, resource-intensive ·
Thousands of crosses evaluated across multiple seasons ·
Marker-Assisted Selection (MAS): DNA markers identify resistance genes early ·
Teagasc: 2,000+ varieties with MAS, pedigree & phenotypic trial data · AI optimisation enables strategic crossing decisions

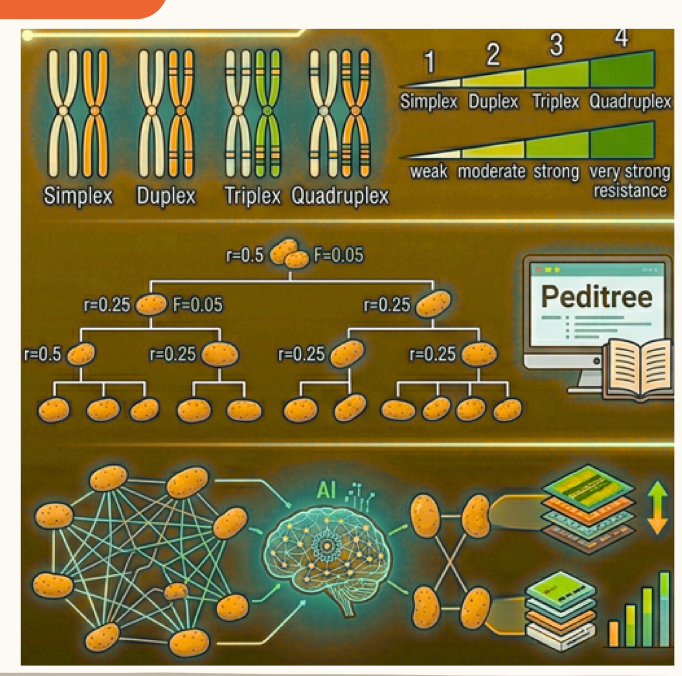


RESEARCH OBJECTIVE

AI-based decision-support system integrating MAS marker data, pedigree-based kinship & phenotypic trial data · Maximise disease resistance stacking · Maintain agronomic performance · Control inbreeding

BACKGROUND AND SIGNIFICANCE

Autotetraploid crop — 4 chromosome copies ·
Peditree quantifies relatedness between breeding lines ·
Genomic markers + pedigree data → efficient resistance gene stacking

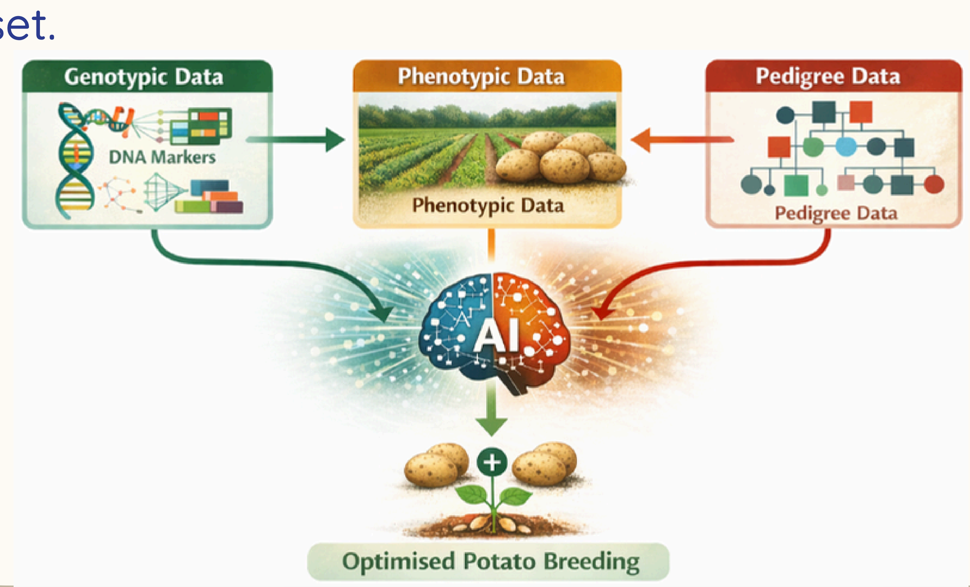


THE DATASET (TEAGASC - CONFIDENTIAL)

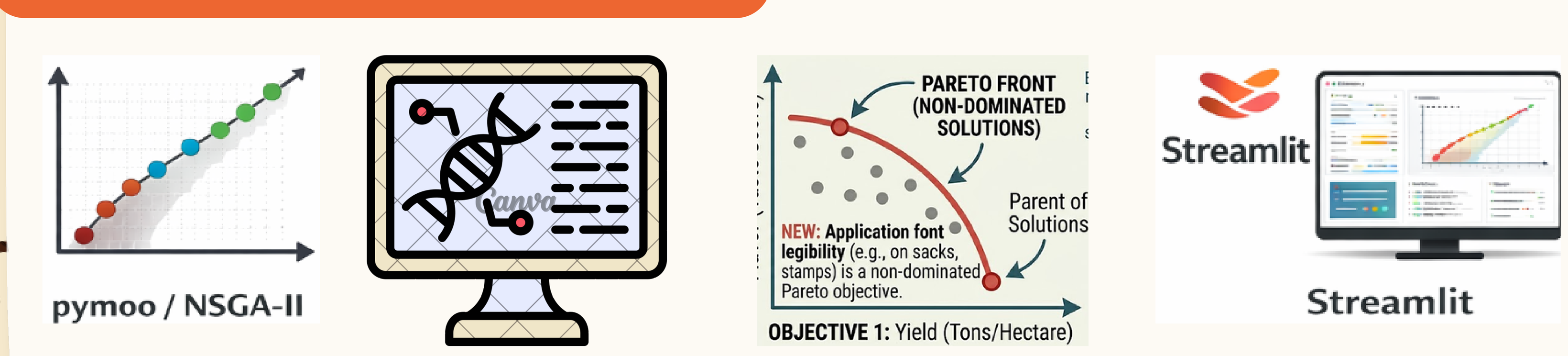
pedigree records from Teagasc's active potato breeding programme — 2,323 varieties.
Allele dosage encoding: 0 = nulliplex (no alleles), 1 = simplex, 2 = duplex, 3 = triplex, 4 = quadruplex (fully resistant)

METHODOLOGY

Dosage encoding 0–4; BLUP normalisation across multi-year trial environments.
Apply Genetic Algorithm using single objective.
NSGA-II Cross Optimisation Each candidate cross encoded as (female, male) — an integer pair indexing the variety dataset.



TECHNOLOGIES



EARLY INDICATIONS & NEXT STEPS

Early Indications:

- Potato encoding 0–4; BLUP normalisation across multi-year trial environments.
- The single-objective GA is also performing strongly, with resistance stacking scores.

Next steps include:

- Expand the model into a full NSGA-II framework using all three objectives.
- Generate a Pareto front of optimal crossing pairs.
- Develop a Streamlit-based explainable prototype to display ranked crosses and marker-wise resistance profiles.
- Compare results with a random baseline using the hypervolume indicator

REFERENCES

Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), pp. 182–197. doi: 10.1109/4235.996017.
Endelman, J. B. (2025). Genomic prediction of heterosis, inbreeding control, and mate allocation in outbred diploid and tetraploid populations. Genetics, 229(2), p. iyae193. doi: 10.1093/genetics/iyae193.
Endelman, J. B., Carley, C. A. S., Bethke, P. C., Coombs, J. J., Clough, M. E., da Silva, W. L., De Jong, W. S., Douches, D. S., Haynes, K. G., Holm, D. G., Miller, J. C., Novy, R. G., Palta, J. P., Parish, D. L., Porter, G. A., Sathuvalli, V. R., Thompson, A. L. and Yencho, G. C. (2018). Genetic variance partitioning and genome-wide prediction with tetraploid inheritance. Genetics, 209(1), pp. 77–87. doi: 10.1534/genetics.118.300685.
Liu, H. et al. (2025). Deep learning for multi-trait genomic prediction in autotetraploids. Plant Methods, [In Press/Online First]. doi: 10.1186/s13007-024-01321-0.
Mousavi, S. M. N. et al. (2024). Heuristic optimization and mate allocation in modern plant breeding. Frontiers in Plant Science, 15, p. 1361894. doi: 10.3389/fpls.2024.1361894.
Slater, A. T., Cogan, N. O. I., Hayes, B. J., Nguyen, T. G., Daetwyler, H. D. and Forster, J. W. (2016). Improving genetic gain in the Australian potato breeding program with genomic selection. Frontiers in Plant Science, 7, p. 1103. doi: 10.3389/fpls.2016.01103.
Van Berloo, R. and Hutten, R. C. B. (2005). Peditree: pedigree database analysis and visualization for breeding and science. Journal of Heredity, 96(6), pp. 720–724. doi: 10.1093/jhered/esi104.
Zhu, J. et al. (2025). Advances in genetics and plant breeding with artificial intelligence and data science. International Journal of Agriculture, 7(9), pp. 772–780.