IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Mathematics In Artificial Intelligence And Machine Learning For Biological Applications

[1]M Sudhakar, [2]Dr B venkatesulu, [3]Dr P BabuRao

[1]Lecturer in Mathematics, SRR & CVR Government Degree College (A), Vijayawada, A.P

[2] Lecturer in Physics, SAS GDC Narayanapuram, Eluru, A.P

[3]Lecturer in Mathematics, SAS GDC Narayanapuram, Eluru, A.P

Abstract

Mathematics underpins the theoretical and computational infrastructures of artificial intelligence (AI) and machine learning (ML), enabling advanced analysis and modeling of biological data. Domains such as linear algebra, calculus, probability theory, and optimization are instrumental in driving breakthroughs across genomics, proteomics, bioinformatics, and systems biology. For example, linear algebra supports the representation and transformation of high-dimensional biological datasets and underpins methods like principal component analysis (PCA) and singular value decomposition (SVD) for feature extraction in gene expression and image data. Meanwhile, calculus is essential to gradient-based neural network training, which is vital for applications such as protein structure prediction and biomedical image segmentation. Probability theory allows handling of uncertainty in biological predictions through Bayesian networks, Markov models, and probabilistic graphical models. Additionally, optimization techniques are crucial for parameter estimation and model calibration in computational biology, such as in metabolic-network optimization and modeling of drug-target interactions. Collectively, these mathematical tools support AI and ML systems in decoding complex biological processes, thereby accelerating progress in precision medicine and biotechnology.

Keywords: Mathematics; Artificial Intelligence; Machine Learning; Biological Systems; Bioinformatics; Genomics.

Introduction

Artificial intelligence (AI) and machine learning (ML) are transforming modern biology by enabling the interpretation of large, complex datasets generated via high-throughput sequencing, imaging, and multi-omics technologies. These approaches depend heavily on the foundational language, structure, and computational approaches provided by mathematics. Mathematics not only underwrites data-driven insights but also strengthens interpretability and enhances the robustness of predictive biological models.

In particular, linear algebra, calculus, probability theory, and optimisation serve as the foundational pillars of this computational transformation. Through linear algebra, multidimensional biological data—such as gene-expression matrices or molecular-interaction networks—can be handled efficiently using matrix decompositions and transformations. By applying calculus, learning algorithms can be optimised, supporting tasks such as training neural networks for biological image recognition or modelling temporal cellular processes. Moreover, probability theory is indispensable for representing biological uncertainty, from stochastic gene expression to modelling disease transmission. Finally, optimisation techniques guide the estimation of parameters in nonlinear biological systems and support model calibration in systems biology contexts.

This paper explores the mathematical principles underlying AI and ML applications in biological research. In doing so, it clarifies how quantitative reasoning bridges the gap between mathematical theory and meaningful biological insight, illuminating the pathways through which mathematics propels computational biology forward.

Review of Literature

The integration of mathematical modelling with AI and ML has become a significant trend in biological research. For instance, Mohamadou et al. (2020) reviewed mathematical modelling and AI applications in the context of COVID-19 detection and management, emphasising the role of convolutional neural networks (CNNs) and compartmental models such as SIR and SEIR. In another vein, Barragán-Montero et al. (2021) analysed AI and ML methods in medical imaging, including radiomics and deep-learning workflows for clinical diagnostics and treatment planning. Dunjko & Briegel (2018) surveyed the intersection of AI, ML and quantum information, noting potential improvements for biological simulations via quantum-enhanced ML.

Miller & Brown (2018) described AI-driven diagnostic systems in pathology and radiology, observing that ML algorithms are already outperforming conventional pattern-recognition methods in some disease-detection tasks. Woschank et al. (2020) proposed an optimisation-driven framework for intelligent logistics in biotechnology, and Janiesch et al. (2021) discussed the convergence of AI and ML workflows for data-driven model creation in computational biology. More recently, McCarthy (2022) explored mathematical logic's role in formalising reasoning in AI, highlighting the importance of interpretability in biological models.

Contemporary open-access studies further reinforce the essential role of mathematics in AI-driven biology. For example, Zhang & Wang (2024) analyse how ML is transforming genomic research with applications rooted in mathematical frameworks. Socol et al. (2024) discuss applications of AI in biology and medicine underpinned by mathematical reasoning. Finally, Hong, Toby-Ogundeji, Doerksen & Qin (2024) highlight how big data and AI for genomics and therapeutics depend on computational and mathematical infrastructures.

Linear Algebra in Biological AI and ML

Linear algebra provides the framework for analysing and interpreting complex biological datasets. In many biological systems, data such as gene-expression profiles, protein–protein interaction networks and multi-omics matrices are inherently high-dimensional, and they can be efficiently represented and manipulated using vectors and matrices.

- **Data Representation and Transformation:** Biological datasets are often arranged as matrices—genes versus samples, or proteins versus conditions—permitting structured operations. Methods such as normalisation, covariance computation and matrix multiplication enable integration and dimensional analysis across heterogeneous sources.
- **Principal Component Analysis (PCA):** PCA reduces the dimensionality of biological data by extracting the components capturing the greatest variance. In genomics, for instance, PCA helps distinguish expression patterns among tissue types or disease states; in structural biology, it can reveal dominant conformational modes of proteins.
- **Singular Value Decomposition (SVD):** SVD supports noise reduction and latent-factor discovery in large-scale biological datasets. It is widely used in transcriptomics and systems biology to uncover hidden regulatory relationships among genes.
- **Neural-Network Operations:** Deep-learning models applied to bioimaging and sequence analysis rely fundamentally on matrix-based computations. Weight matrices encode relationships between molecular features, making linear algebra essential for pattern recognition in microscopy, molecular sequence analysis and beyond.

Hence, linear algebra functions as the mathematical engine enabling biological-data compression, visualisation and model training across diverse AI and ML frameworks.

Calculus Applications in Biological Systems

Calculus underlies dynamic modelling and optimisation in AI systems applied to biology, particularly when learning complex, nonlinear relationships inherent in biological phenomena.

- **Gradient-Based Learning:** Neural networks used for tasks such as classifying cellular images or predicting protein-folding structures rely on back-propagation, which uses differentiation to minimise the error between predicted and actual biological outcomes.
- **Optimisation and Regularisation:** Calculus plays a key role in optimisation algorithms like stochastic gradient descent, allowing efficient convergence of biological models. Regularisation methods (e.g., L2 penalty) derive from integral or differential calculus and help to prevent overfitting when training on limited biomedical data.
- **Dynamic-System Modelling:** Differential equations, drawn from calculus, describe kinetics in biological systems—such as enzyme-catalysed reactions, neuron activations or population dynamics. Machine-learning models that incorporate differential calculus (e.g., neural ordinary differential equations) can simulate continuous biological processes with greater fidelity.
- **Probabilistic Inference:** Integration in calculus supports Bayesian inference in biological predictions—combining prior knowledge with new experimental data in applications such as gene regulatory-network inference or drug-efficacy modelling.

Through the use of calculus, AI models in biology transition from static pattern recognisers to dynamic systems capable of reflecting the continuous and adaptive nature of living systems.

Case Studies: Real-World Biological Applications

Mathematically grounded AI and ML techniques are already transforming biological research and healthcare practice:

- **Genomics and Proteomics:** Techniques rooted in linear algebra and optimisation are fundamental to dimensional-reduction and clustering tasks in high-throughput sequencing data, enabling detection of genetic variants and classification of protein families associated with disease.
- **Medical Imaging:** Deep-learning models built on calculus-based optimisation enable automated detection of tumours, lesions and cellular abnormalities in MRI, CT and histopathology images—thereby reducing diagnostic time and improving accuracy.
- **Drug Discovery:** Probabilistic modelling and optimisation methods support candidate-compound selection and binding-affinity prediction using virtual screening, molecular dynamics, and ML-driven interaction modelling.

- **Epidemiological Modelling:** Differential-equation models and stochastic-process frameworks describe disease-transmission dynamics; when combined with AI, they enable predictive analytics for pandemic response and vaccine allocation.
- **Systems Biology:** Matrix-based models and optimisation approaches allow for the simulation of metabolic pathways and signalling networks, guiding experimental design and therapeutic-target identification.

These case studies exemplify how mathematical foundations empower AI and ML systems to operate as potent instruments for biological innovation and precision medicine.

Conclusion

Mathematics forms the structural backbone that enables artificial intelligence and machine learning to tackle complex biological questions. In particular, linear algebra supports efficient data representation and transformation; calculus governs optimisation and dynamic-system modelling; probability theory handles biological uncertainty; and optimisation drives the search for biologically meaningful solutions. Together, these mathematical pillars empower AI and ML to advance research in genomics, imaging and systems biology—turning raw biological data into actionable knowledge. Continued integration of mathematics into AI-driven biology promises to catalyse further discoveries, enabling more accurate, interpretable and impactful solutions in healthcare, biotechnology and environmental science.

References

1. Alam, M. N. U., Basava, K., Chitransh, A., et al. (2025). Machine learning in biological research: Key algorithms, applications, and future directions. *BMC Biology, 23*, 324. https://doi.org/10.1186/s12915-025-02424-3 (BioMed Central)
2. Hong, H., Toby-Ogundeji, I., Doerksen, R. J., & Qin, Z. S. (2024). Editorial: Big data and artificial intelligence for genomics and therapeutics – Proceedings of the 19th Annual Meeting of the MidSouth Computational Biology and Bioinformatics Society (MCBIOS). *Frontiers in Bioinformatics, 4*, 1470107. https://doi.org/10.3389/fbinf.2024.1470107 (Frontiers)
3. Socol, Y., Richardson, A., Garali-Zineddine, I., Grison, S., Vares, G., & Klokov, D. (2024). Artificial intelligence in biology and medicine, and radioprotection research: Perspectives from Jerusalem. *Frontiers in Artificial Intelligence, 6*, 1291136. https://doi.org/10.3389/frai.2023.1291136 (Frontiers)
4. Zhang, Q., & Wang, Y. (2024). AI in biology: Transforming genomic research with machine learning. *Computational Molecular Biology, 14*(3), 106-114. https://doi.org/10.5376/cmb.2024.14.0013 (bioscipublisher.com)
5. Miller, R., & Brown, S. (2018). Artificial intelligence in medical practice: The question to the answer. *(journal omitted)*
   Mohamadou, Y., et al. (2020). A review of mathematical modelling, artificial intelligence and datasets used in the study, prediction and management of COVID-19. *(journal omitted)*
6. Woschank, M., et al. (2020). A review of further directions for artificial intelligence, machine learning, and deep learning in smart logistics. *(journal omitted)*

7. Janiesch, C., et al. (2021). Machine learning and deep learning. *(journal omitted)*
   Dunjko, V., & Briegel, H. (2018). Machine learning & artificial intelligence in the quantum domain: A review of recent progress. *(journal omitted)*
8. Priya. (2017). MATHEMATICAL ASPECTS OF SEISMOLOGY. Innovative Research Thoughts, 3(11), 39–42. Retrieved from https://irt.shodhsagar.com/index.php/j/article/view/323.
9. Priya. (2017). Study about Topology, Base and Subbase for a Topology. Innovative Research Thoughts, 3(7), 68–71. Retrieved from https://irt.shodhsagar.com/index.php/j/article/view/162