

Modeling plant stress responses: integrating gene expression dynamics and omics data with Physics-Informed Neural Networks

Supervision

Lorenzo Sala from team Dynenvie, MaIAGE, INRAE

Silvia Bottini from team SMILE, Institut Sophia Agrobiotech, INRAE

Contact

Lorenzo Sala, lorenzo.sala@inrae.fr

Silvia Bottini, silvia.bottini@inrae.fr

To apply send an email with CV and motivation letter with subject 'APPLICATION FOR PINNSVAE INTERNSHIP'.

Scientific context

This project focuses on developing advanced AI methodologies to model and understand dynamic biological systems, particularly plant-pathogen interactions. The primary goal is to leverage Physics-Informed Neural Networks (PINNs) and Variational Autoencoders (VAEs) to improve the integration and interpretation of omics data in these systems. Specifically, the internship will involve building an initial model based on Ordinary Differential Equations (ODEs) to describe the gene expression dynamics of plants, which will be integrated into the existing HIVE platform [1].

HIVE (Horizontal Integration analysis using VAE), developed by Dr. Bottini's team, is a framework for analyzing transcriptomics data from unpaired experiments. The internship will begin with designing an ODE system to model the temporal changes in gene expression under external stresses such as pathogen attacks or environmental conditions [3]. This model will simulate the evolution of gene counts as longitudinal data, enabling the characterization and inference of key interaction parameters. By identifying molecular signatures and their activation timings during stress responses, the project seeks to uncover the defense mechanisms that plants deploy during infection.

Once the ODE model is established, it will be extended into a PINN framework [2] to estimate parameters efficiently. This involves constructing a loss function that incorporates ODE constraints, ensuring the model adheres to biological principles. The neural network will be trained on synthetic data to fit gene expression dynamics and infer the interaction parameters. A well-adapted neural network architecture will be developed to balance predictive accuracy with interpretability, capturing the key dynamics of the biological system.

The intern will focus on integrating the ODE model into the HIVE platform, facilitating the identification of critical data features and capturing plant-pathogen interaction dynamics in a reduced-dimensional space. While full integration and application to real-world data may extend beyond the internship, the project will provide a solid foundation in modeling dynamic biological systems using state-of-the-art AI techniques.

Internship objectives

The primary goal of this internship is to develop a dynamic model of gene expression under stress conditions, leveraging advanced AI techniques to integrate biological constraints and omics data.

The objectives of the internship include:

1. **Developing an ODE-based model:** Construct an ODE system to represent the temporal dynamics of plant gene expression under external stressors, such as pathogen infections or environmental conditions. The model will simulate longitudinal gene expression data, enabling the identification of key molecular interactions.
2. **Extending the model to a PINN framework:** Transform the ODE model into a PINN to provide an alternative solver that is more efficient in case of an extensive number of genes. Develop a loss function that aligns the PINN predictions with the ODE system, ensuring the model adheres to biological principles.
3. **Training and validation using synthetic data:** Train the PINN model on synthetic datasets to validate its ability to infer interaction parameters and predict gene expression dynamics. This step will establish the approach’s feasibility and pave the way for integration with real-world data.
4. **Contributing to the HIVE platform:** Begin integrating the developed ODE/PINN models into the HIVE platform to enhance its capability in analyzing transcriptomics data and capturing the dynamics of plant-pathogen interactions in a reduced-dimensional space.

By achieving these objectives, the internship will provide a deeper understanding of plant responses to stress at the molecular level and contribute to advancing AI-driven methodologies in biological modeling.

Internship outline

The candidate will benefit from the expertise of MaIAGE’s modeling team and the biocomputational insights of the SMILE team at Institut Sophia Agrobiotech, who developed the HIVE platform and were responsible for pre-processing the data.

- Profile: M2 students in applied mathematics or related fields.
- Duration: 6 months (flexible, ideally starting from March 2025).
- Placement location: The intern will be based at MaIAGE, at the INRAE center of Jouy-en-Josas (78).
- Funding: This internship is supported by DATAIA convergence institute as part of the "Programme d’Investissement d’Avenir", (ANR- 17-CONV-0003) operated by INRAE
- Internship Allowance: ~ 650 EUR / month.

Candidate skills

The candidate should possess solid skills in modeling, ODEs, and neural networks.

Proficiency in Python scientific programming is required.

We seek a candidate with a strong interest in working at an interdisciplinary level between mathematics and biology.

References

- [1] G. Calia, S. Marguerit, A. P. Zotta Mota, M. Vidal, M. Seynabou-Fall, H. T. Nguyen, A. Bhat, H. Schuler, C. Gwizdek, A. C. M. Brasileiro, et al. Disentangling plant response to biotic and abiotic stress using hive, a novel tool to perform unpaired multi-transcriptomics integration. *bioRxiv*, pages 2024–03, 2024.
- [2] P. J. Hossie, B. Laroche, T. Malou, L. Perrin, T. Saigre, and L. Sala. Simulating interactions in microbial communities through physics informed neural networks: towards interaction estimation. 2024.
- [3] X. Liu, D. Igarashi, R. A. Hillmer, T. Stoddard, Y. Lu, K. Tsuda, C. L. Myers, and F. Katagiri. Decomposition of dynamic transcriptomic responses during effector-triggered immunity reveals conserved responses in two distinct plant cell populations. *Plant Communications*, 2024.