Politecnico di Torino

Dipartimento di Automatica e Informatica

PhD in Computer and Control Engineering

37° cycle

Supervisor

Daniele Apiletti

Theory-guided Data Science models

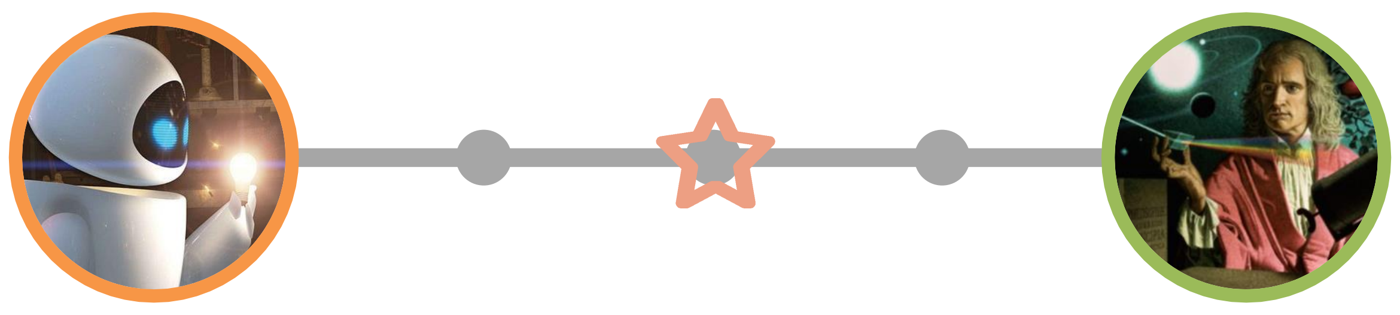
PhD Candidate: *Simone MONACO*

Email: simone.monaco@polito.it

1. Context

Modern deep-learning models require an ever-increasing amount of data. **Can we mitigate their data and power hungriness by injecting prior knowledge?**

While theory-based models are standard in many fields but rarely take advantage of the benefits of data science techniques, machine learning models are often purely data-driven, losing the opportunity to exploit the knowledge behind the data themselves. This poster presents a few of our works combining theory and data science.



2. Challenges and Objectives

Challenges

(i) Limited generalizability across domains and

(ii) Complex construction of structured knowledge (especially in non-physical phenomena).

Objectives

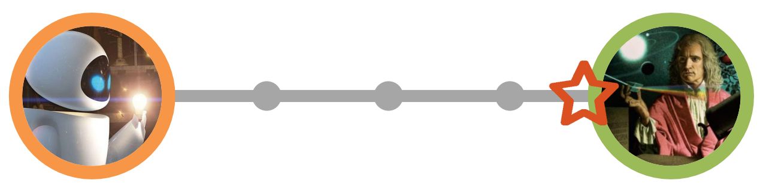
(i) Improving efficiency by knowledge integration

(ii) Enhancing model convergence speed and robustness

(iii) Promoting knowledge generalization and heterogenous knowledge integration

3. Methods and Results

PINNs for differential equations (no data at all)



Physics-informed Neural Networks (PINNs) are designed for solving partial differential equations without external data, relying only on domain dimensions. Networks are trained to align the output with the equation and boundary conditions, using automatic differentiation to obtain differential forms. Direct application of this approach can, however, be very ineffective.

In our work[1], we provide a cross-domain experimental evaluation of state-of-the-art strategies, designed to overcome such limitations. Experimental results highlight strengths and pitfalls of current techniques and show that currently no strategy can successfully generalize Fig. 1.

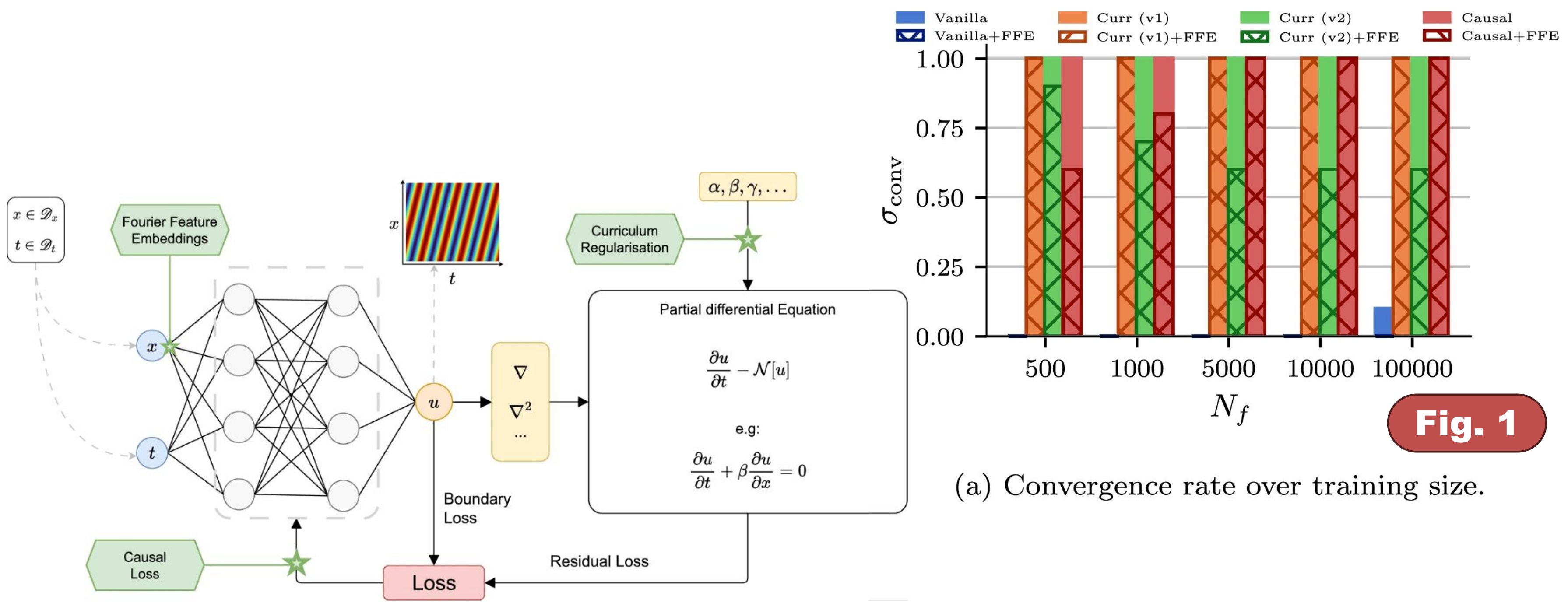
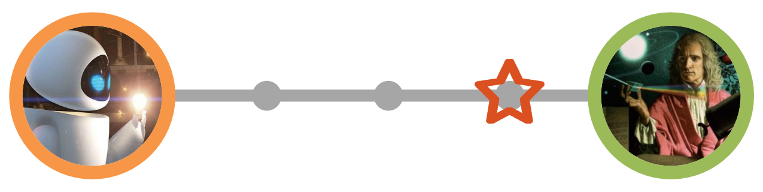


Fig. 1

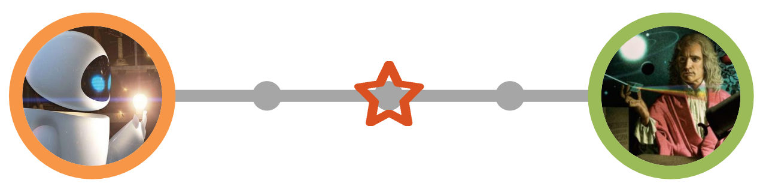
(a) Convergence rate over training size.

Enforcing Hard constraints between inputs/outputs (less need for data)



In the presence of some input-output relations, employing hard analytical-constrained networks (ACNet) can shrink the output search space. Analogous results can be obtained softly by adding an unsupervised loss function (LCNet). By comparing these models with traditional neural networks, we observed that domain-knowledge injection effectively improves network performance and robustness with fewer training data [2] Fig. 2.

Symmetrical invariances (adding properties with extra knowledge)



We designed a data augmentation strategy applied to a baseline model in jet tagging (ParticleNet) with a variable strength to enforce the model understanding of the Lorentz invariance principle[3]. Our approach could enhance the model to respect this additional invariant property, without adding complexity, and without affecting performance Fig. 3.

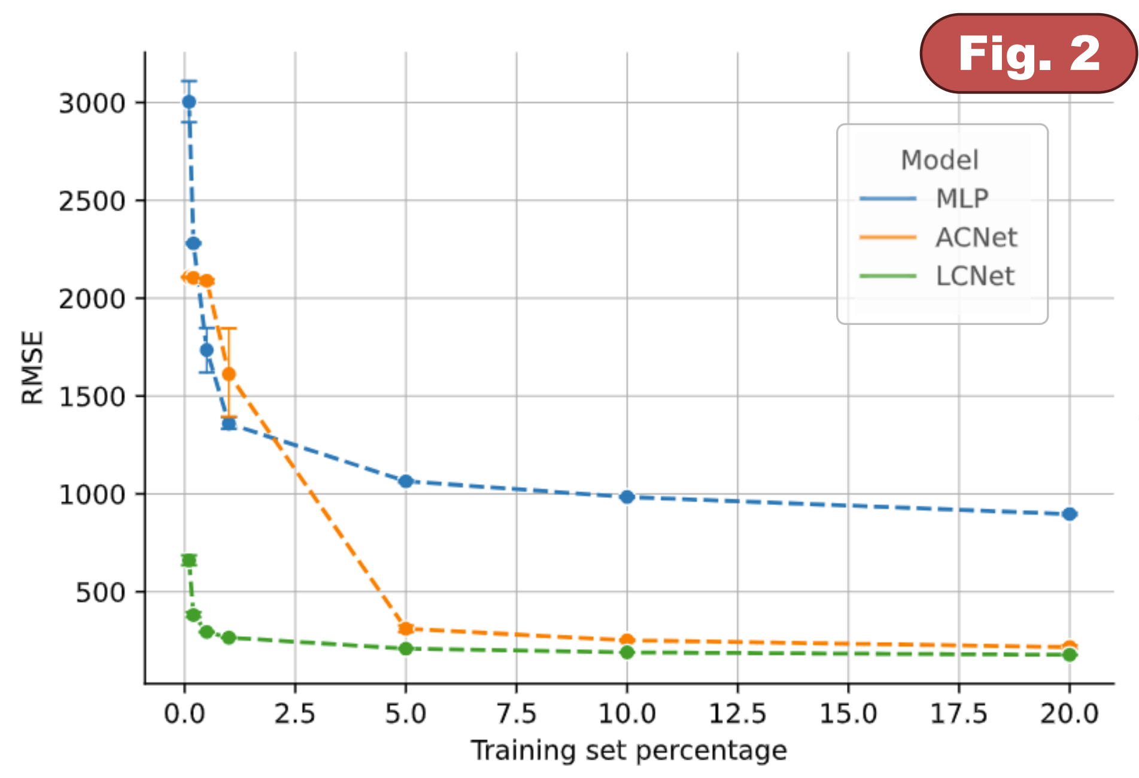


Fig. 2

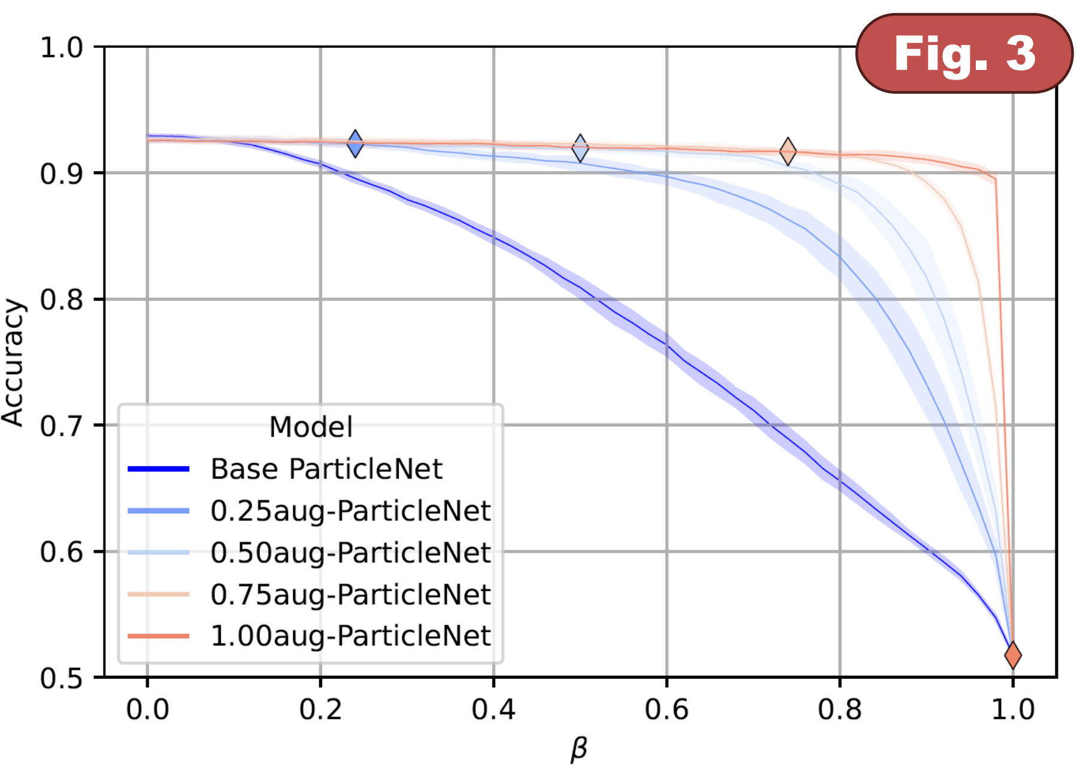


Fig. 3

4. Conclusions

Leveraging domain knowledge is crucial for enhancing performance and resilience of data-driven models in diverse applications. The various forms of knowledge integration may yield different impacts on model architectures. This insight highlights the key challenge in these approaches: achieving generalizability.

In the future, we aim to create more generalized knowledge injection frameworks, exploring patterns to determine the benefits of each approach. Additionally, we plan to explore innovative methods for incorporating unstructured knowledge from multiple domains.

5. References

1. Monaco, Simone, and Daniele Apiletti. "Training physics-informed neural networks: One learning to rule them all?." Results in Engineering 18 (2023).
2. Monaco, Simone, Daniele Apiletti, and Giovanni Malnati. "Theory-Guided Deep Learning Algorithms: An Experimental Evaluation." Electronics 11.18 (2022).
3. Monaco, Simone, Sebastiano Barresi, and Daniele Apiletti. "Lorentz-invariant augmentation for high-energy physics." Communications in Computer and Information Science. Springer, 2023.