ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
UNIVERSITY OF PIRAEUS

ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Ω.Π.Μ.Σ. ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

ΔΗΜΟΚΡΙΤΟΣ
DEMOKRITOS

ΙΝΣΤΙΤΟΥΤΟ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

Προτεινόμενα Θέματα Διπλωματικής Εργασίας

Από

Διδάσκοντα Χριστόφορο Ρεκατσίνα

A. Shallow Machine Learning Architectures for Enhancing Computational Modelling in Physics and Engineering

Qualifications required: Python programming; Machine Learning; basic knowledge of numerical methods and scientific computing.

This thesis proposes a data-efficient machine learning framework for enhancing computational modelling in physics and engineering, with applications ranging from quantum mechanics and molecular dynamics to nonlinear mechanics. The work focuses on three main methodologies: Physics-Informed Neural Networks (PINNs), Optimization of a Discrete Loss (ODIL), and Neural Operators.

PINNs will be used to embed governing physical equations directly into the learning process, allowing the model to solve forward and inverse problems with limited data. This is particularly relevant for problems governed by differential equations, such as quantum wave equations, transport phenomena, and nonlinear mechanical response.

ODIL will be investigated as an alternative physics-based learning strategy, where the governing equations are discretized first and the loss function is constructed directly from the residuals of the discrete numerical problem. This approach can improve numerical stability and provide a closer link between machine learning and classical computational solvers.

Neural Operators will be explored for learning mappings between functions, such as material parameters, boundary conditions, forcing terms, or potentials, and the corresponding physical response. This makes them suitable for accelerating repeated simulations in molecular dynamics, quantum mechanics, and nonlinear multiphysics problems.

The thesis will develop and compare these approaches on representative benchmark problems, starting from simplified quantum or molecular systems and extending toward nonlinear mechanics applications. Emphasis will be placed on data efficiency, physical consistency, computational acceleration, and the ability of the models to generalize across

different conditions. The final objective is to assess how physics-informed and operator-learning methods can support faster and more reliable computational workflows for complex physics and engineering problems.

Supervisors: Christoforos Rekatsinas, George Giannakopoulos, Panagiotis Krokidas

References

Physics-Informed Neural Networks without Loss Balancing: A Direct Term Scaling Approach for Nonlinear 1D Problems, https://doi.org/10.12688/f1000research.169129.2

Fourier Neural Operators Explained: A Practical perspective, https://arxiv.org/pdf/2512.01421

Solving inverse problems in physics by optimizing a discrete loss: Fast and accurate learning without neural networks, https://doi.org/10.1093/pnasnexus/pgae005

B. Reinforcement Learning for the Inverse Design of MOF-Integrated Composite Pressure Vessels for Hydrogen Storage

This thesis will investigate the use of **multi agent reinforcement learning for the inverse design of MOF-integrated composite pressure vessels for hydrogen storage**. Inspired by recent work where airfoil optimization was formulated as a Markov Decision Process, the thesis will transfer the same logic to pressure-vessel design. Instead of modifying aerodynamic shapes, the reinforcement learning agent will optimize design parameters such as vessel geometry, wall thickness, composite layup, internal volume allocation, and MOF loading configuration.

The objective is to develop a computational framework that balances **hydrogen-storage efficiency, structural safety, and lightweight design**. The pressure vessel will be evaluated through simplified analytical or numerical models estimating mass, stress levels, deformation, and safety factors under internal pressure. At the same time, MOF-related descriptors, such as porosity, density, surface area, and adsorption capacity, will be used to estimate the contribution of the adsorbent material to hydrogen storage.

The main outcome will be a prototype AI-driven inverse design environment where an agent learns to propose improved pressure-vessel/MOF configurations under engineering constraints. The thesis will demonstrate how reinforcement learning can support the

early-stage design of advanced hydrogen-storage systems by combining structural mechanics, materials selection, and data-driven optimization.

A reinforcement learning approach to airfoil shape optimization, *https://doi.org/10.1038/s41598-023-36560-z*

C. Predicting relevance in scientific information retrieval

Qualifications required: Python programming; Machine Learning and Deep Learning algorithms.

Qualifications desired: Machine learning and Deep learning toolkits (e.g. PyTorch); Natural Language Processing (NLP)
Supervisors: George Giannakopoulos, Artemis Dampa
Contact: ggianna [at] iit.demokritos.gr
Related groups: INSANE Group

Description: The objective of this thesis is to propose and implement methods to improve relevance prediction in scientific information retrieval. During scientific research and literature reviews, researchers often retrieve a large number of documents, datasets, and other resources that need to be managed in order to extract valuable, domain-specific knowledge. To address this challenge, this work will primarily focus on, but is not limited to, the effective representation of scientific documents. This can be achieved by incorporating prior knowledge to better capture the complex, specialized nature of scientific texts. Additionally, the thesis will explore unbiased methods for corpora collection, annotation, and relevance evaluation, as well as relevance classification techniques that integrate both explicit and implicit user feedback. Lastly, query expansion using natural language processing methods will be utilized to automatically expand search queries with related terms, synonyms, and conceptually relevant keywords. Ultimately, the aim is to develop a more accurate and efficient system for retrieving and evaluating scientific literature, enhancing researchers' ability to access the most relevant information.

Cohan, Arman, et al. "Specter: Document-level representation learning using citation-informed transformers."

D. Explainable Physics-Informed Machine Learning: A Review and Comparative Assessment of Methods

Qualifications required: Python programming; Machine Learning and Deep Learning algorithms

Supervisors: Christoforos Rekatsinas, Vasilis Gkatsis
Contact: crek [at] iit.demokritos.gr
Related groups: INSANE Group

Abstract

Modern machine learning methods have shown strong potential in physics and engineering problems, but their use in scientific domains is often limited by data scarcity, poor interpretability, and weak guarantees of physical consistency. Physics-informed machine learning addresses these limitations by embedding prior knowledge into the learning process, including governing equations, conservation laws, boundary conditions, symmetry constraints, expert rules, or simulation-based knowledge. Although these models are often presented as more interpretable than purely data-driven approaches, the degree and usefulness of their explainability remain open questions.

This thesis will focus mainly on a structured review of explainable physics-informed machine learning methods. The student will study and classify approaches such as Physics-Informed Neural Networks, Physics-Guided Neural Networks, Neural Operators with physical constraints, hybrid physics–ML models, symbolic regression, sparse identification of governing equations, and neuro-symbolic methods. The review will cover approximately 80–90 scientific references and will organize the literature according to the type of physical knowledge used, the target application, the level of interpretability, and the explainability technique employed.

A limited implementation component will also be included. The student will reproduce or develop simple benchmark examples to compare selected methods in terms of accuracy, data efficiency, physical consistency, and explainability. Possible test cases may include simple dynamical systems, diffusion-type equations, or low-dimensional mechanics problems. The objective is not to develop a new large-scale model, but to provide a critical and well-structured assessment of how explainable physics-informed machine learning methods can be evaluated and compared.

References

Discussing the spectrum of physics-enhanced machine learning: a survey on structural mechanics applications, https://doi.org/10.1017/dce.2024.33

Structural identification with physics-informed neural ordinary differential equations, https://doi.org/10.1016/j.jsv.2021.116196

The language of hyperelastic materials, https://doi.org/10.1016/j.cma.2024.117053

E. Grammar-Based Generative Modeling for architected materials with Property Optimization

Qualifications required: Python programming; Machine Learning and Deep Learning algorithms

Supervisors: Christoforos Rekatsinas, Vasilis Vasilis Sioros, Panagiotis Krokidas
Contact: crek[at] iit.demokritos.gr
Related groups: INSANE Group

Abstract

Metal-Organic Frameworks (MOFs), and architected composite materials, hold significant promise for advanced industrial applications, including gas separation, H2 adsorption, enhanced ductility, and damping. Despite progress in predictive methods, challenges persist in designing both MOFs and architected composites with superior functional properties, such as high adsorption capacity, mechanical resilience, and energy dissipation. While deep learning (DL)-based generative models have shown potential for molecular and material design, their reliance on large datasets limits their applicability in data-scarce domains. Grammar-based generative approaches offer an interpretable and data-efficient alternative, enabling the explicit integration of chemical and structural constraints. This project aims to adapt a grammar-based generative method for the design of MOFs and architected composite materials, incorporating property prediction models to optimize structures for targeted applications, including gas separation, hydrogen storage, mechanical durability, and damping performance.

References

Data-Efficient Graph Grammar Learning for Molecular Generation

A comprehensive transformer-based approach for high-accuracy gas adsorption predictions in metal-organic frameworks

Bio-inspired discontinuous composite materials with a machine learning optimized architecture

Appendix

Advancing Bayesian Optimization: Efficient Strategies for Accelerated Materials Discovery

Qualifications required: Python programming; Machine Learning

Qualifications desired: Machine learning and Deep learning toolkits (e.g. PyTorch); genetic algorithms, bayesian optimization, active learning

Supervisors: George Giannakopoulos, Panagiotis Krokidas, Christoforos Rekatsinas
Contact: ggianna [at] iit.demokritos.gr, p.krokidas [at] iit.demokritos.gr
Related groups: INSANE Group

Description: Bayesian Optimization (BO) is an AI-driven technique designed to efficiently explore and sample large design spaces, particularly in domains where data are sparse, experiments are costly, and simulations are computationally intensive. This approach is especially valuable in the development of novel materials, enabling the discovery of optimized materials with enhanced performance while significantly reducing the number of physical or computational experiments required. However, traditional BO methodologies are constrained by high memory demands, largely due to the reliance on Gaussian Processes (GPs) as surrogate models. This limitation leads to a computational complexity that scales as $O(n^3)$, making it impractical for large-scale or high-dimensional problems. To address these challenges, next-generation BO algorithms must incorporate more efficient strategies. This thesis will focus on investigating advanced approaches to improve the efficiency of BO, such as batch sampling techniques, dynamically optimized batch sizes, adaptive exploration-exploitation trade-offs, efficient surrogate models, and multi-fidelity methods inspired by state-of-the-art advancements. The goal is to develop and evaluate innovative methods that maintain the predictive power of BO while significantly reducing memory and computational overhead, ultimately enabling its broader applicability in materials science and other high-impact fields.

References

Siemenn et al. Fast Bayesian optimization of Needle-in-a-Haystack problems using zooming memory-based initialization (ZoMBI), npj Computational Materials, 2023

Gantzler et al. Multi-fidelity Bayesian optimization of covalent organic frameworks for xenon/krypton separations, Digital Discovery, 2023