ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Informed Machine Learning for Complex Data

Luca Oneto[1], Nicolò Navarin[2], Alessio Micheli[3], Luca Pasa[2],
Claudio Gallicchio[3], Davide Bacciu[3], Davide Anguita[1] *

1 - University of Genoa - Via Opera Pia 11a, 16145, Genova, Italy
2 - University of Padua - Via Trieste 63, 35121, Padua - Italy
3 - University of Pisa - Largo Bruno Pontecorvo 3, 56127, Pisa, Italy

Abstract. In the contemporary era of data-driven decision-making, the application of Machine Learning (ML) on complex data (e.g., images, text, sequences, trees, and graphs) has become increasingly pivotal (e.g., Large Language Models and Graph Neural Networks). In this context, there is a gap between purely data-driven models and domain-specific knowledge, requirements, and expertise. In particular, this domain specificity needs to be integrated into the ML models to improve learning generalization, sustainability, trustworthiness, reliability, security, and safety. This additional knowledge can assume different forms, e.g.: software developers require ML to comply with many technical requirements, companies require ML to comply with economic and environmental sustainability, domain experts require ML to be aligned with physical and logical laws, and society requires ML to be aligned with ethical principles. This special session gathers valuable contributions and early findings in the field of Informed ML for Complex Data. Our main objective is to showcase the potential and limitations of new ideas, improvements, or the blending of ML and other research areas in solving real-world problems.

1 Introduction

In the contemporary era of data-driven decision-making, the application of Machine Learning (ML) has emerged as a cornerstone of innovation across various domains [1–3]. The versatility of ML models in handling complex data types such as images, text, sequences, trees, and graphs has opened up new avenues for solving intricate problems [4]. The development and deployment of Large Language Models [2, 5] and Graph Neural Networks [6, 7] exemplify the strides made in leveraging complex data for intelligent decision-making.

However, the journey of ML from theoretical frameworks to practical applications is fraught with challenges. A significant gap exists between purely data-driven models and the specific knowledge, requirements, and expertise inherent in various domains. Bridging this gap is essential for enhancing the generalization, sustainability, trustworthiness, reliability, security, and safety of ML models. Integration of domain-specific knowledge into ML models is not

*This work is partially supported by (i) EU - NGEU National Sustainable Mobility Center (CN00000023) Italian Ministry of University and Research Decree n. 1033—17/06/2022 (Spoke 10), (ii) project SERICS (PE00000014) under the NRRP MUR program funded by the EU - NGEU, (iii) project FAIR (PE00000013) under the NRRP MUR program funded by the EU - NGEU, (iv) EU-EIC EMERGE (Grant No. 101070918) and (v) NEURONE, a project funded by the Italian Ministry of University and Research (PRIN 20229JRTZA).

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

merely a technical requirement, but a multidimensional necessity that encompasses technical, economic, environmental, and ethical aspects [4, 7–21].

The domain-specific integration of knowledge, requirements, and expertise into ML models requires one to address several critical needs. Software developers require ML models that comply with numerous technical specifications and performance standards [4, 7, 8]. This includes considerations for computational efficiency, scalability, and interoperability with existing systems. Companies and industries demand that ML models not only deliver economic value, but also align with sustainable practices [9–11]. This includes optimizing resource utilization, minimizing energy consumption, and reducing the carbon footprint associated with model training and deployment. Domain experts emphasize the need for ML models to align with established physical and logical principles [12–14]. For example, in scientific research and engineering, models must respect conservation laws and other fundamental principles to ensure valid and reliable outputs. The society in general requires ML models to operate within ethical limits, ensuring fairness, transparency, and accountability [15–21]. This is crucial to maintaining public trust and addressing concerns related to bias, privacy, and the societal impact of ML-based technologies.

Recognizing these multifaceted requirements, the special session on *Informed ML for Complex Data* aims to gather valuable contributions and early findings from the field. The primary objective is to showcase the potential and limitations of novel ideas, improvements, and the integration of AI, ML, and other research areas to address real-world problems. The session emphasizes the exploration of innovative approaches that integrate domain-specific knowledge into ML models. A key focus is on demonstrating the practical implications of informed ML models in solving real-world problems. While the integration of domain-specific knowledge presents significant opportunities, it also introduces complexities and limitations. The session provides a platform for discussing these challenges, including the difficulty in accurately encoding domain knowledge, potential biases introduced by domain-specific assumptions, and the computational overhead associated with more complex models.

The integration of domain-specific knowledge into ML models represents a critical evolution in the field of ML. By addressing the gap between data-driven approaches and domain-specific requirements, researchers and practitioners can develop more effective, sustainable, and trustworthy ML solutions.

2 Data-informed

With the vast amount of data that is produced daily, the ability to extract meaningful information from it is becoming increasingly significant. The data tends to show complex relationships, such as the sequence information in a text or the relationships among entities in (graph) databases, social and IoT data. Given the speed with which the data is generated, machine learning models with the ability to directly learn from complex data are becoming more and more popular. Two clear examples are Large Language Models, that are able to exploit

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

textual data to an extent in which they show primitive reasoning abilities [22, 23], or Graph Neural Networks, that are capable of learning task-specific concepts over the extremely versatile graph representation, that can encode as graphs a widespread array of data, from chemical compounds to social networks [24–26]. One of the main challenges in Graph Neural Networks is their applicability to huge datasets, both in terms of computational time and memory requirements [27], also considering continual learning [28]. Graph Neural Networks are the most popular flavour of a family of models commonly known as geometric deep learning [29]. Another example of complex data belonging to the same family and that requires specific deep learning models to deal with are point clouds, that are increasingly popular since thay are generated, for instance, from LiDaR (Light Detection and Ranging) sensors that are becoming more and more common in mobile devices [30].

3 Technically Informed

Integrating ML models into production systems necessitates compliance with a range of technical specifications and performance standards [4, 7, 8, 31].

Efficient models minimize operational costs and latency, enhancing the user experience. Techniques such as model pruning, quantization, and specialized hardware (e.g., GPUs, TPUs) are essential for achieving high computational efficiency. Balancing model complexity and performance is crucial for resource-efficient operations.

Scalable ML models handle varying loads, adapting to increasing data volumes and user demands without performance degradation. Distributed training, model parallelism, and cloud-based services are key techniques. Developers must design horizontally scalable models, leveraging distributed computing environments.

Ensuring ML models seamlessly integrate with existing systems and technologies involves compatibility with various programming languages, frameworks, and data formats. Standardized protocols, APIs, and containerization (e.g., Docker) enhance interoperability, facilitating smoother deployment and integration.

Modern applications require ML models to process complex data like text, graphs, and trees. Specialized techniques, such as Large Language Models for text and Graph Neural Networks are crucial. Developers must ensure models effectively process these data structures for accurate predictions.

The technical debt in ML systems arises from factors like data dependencies, configuration issues, and entanglement, which can significantly degrade system performance over time. The importance of managing these debts through practices such as modular design, testing, and continuous monitoring is crucial calling for more attention to the engineering aspects of ML to ensure sustainable and scalable solutions. It is then required to provide practical recommendations for developers and researchers to mitigate these hidden risks in their ML projects.

Technically Informed ML emphasizes meeting technical standards and per-

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

formance benchmarks, ensuring ML models deliver accurate results and integrate seamlessly into production environments, providing long-term value and sustainability.

4 Environmentally Informed

As the demand for ML continues to grow across various sectors, there is an increasing emphasis on the environmental impact of these technologies [9–11, 32]. Companies and industries now require ML models not only to deliver economic value, but also to align with sustainable practices. This shift towards environmentally informed ML encompasses optimizing resource utilization, minimizing energy consumption, and reducing the carbon footprint associated with model training and deployment.

Effective resource utilization is a critical aspect of environmentally informed ML. This involves using computational resources such as CPUs, GPUs, and memory more efficiently. Techniques like model pruning, which reduces the size of neural networks, and the use of lightweight architectures can significantly lower the computational demands. Moreover, scheduling algorithms that optimize the use of data center resources can lead to more efficient utilization, reducing the environmental impact of running large-scale ML operations.

Reservoir computing [33, 34], a paradigm that takes advantage of the dynamic properties of recurrent neural networks with fixed random connections, offers an energy-efficient alternative for specific ML tasks. Using a fixed random network as the "reservoir", and only training the output layer, reservoir computing significantly reduces computational load and energy consumption [35]. This approach is particularly advantageous for tasks that involve temporal data and real-time processing, where rapid responses with minimal computational overhead are essential.

Energy consumption is a major factor in the environmental footprint of ML models. Training large models, especially deep neural networks, can be extremely energy-intensive. Strategies to minimize energy consumption include adopting more energy-efficient hardware, such as GPUs designed for lower power usage, and optimizing software to reduce unnecessary computations. Additionally, implementing energy-aware algorithms that dynamically adjust the computational load based on energy availability and cost can contribute to significant energy savings.

The carbon footprint of ML models, particularly those that require extensive training on large datasets, can be substantial. An approach to reducing this footprint is through the use of renewable energy sources to power data centers. Companies can also offset their carbon emissions by investing in carbon credits or supporting reforestation projects. Furthermore, it is essential to develop and utilize ML frameworks that prioritize energy efficiency and carbon neutrality. For example, federated learning, which trains models across decentralized devices, can reduce the need to centralize massive amounts of data, thereby reducing overall energy consumption and carbon emissions.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Evaluating the environmental impact of ML models requires a comprehensive lifecycle assessment, from development and training to deployment and maintenance. This includes evaluating the energy and resource requirements at each stage and identifying opportunities to reduce the environmental footprint. For instance, by reusing pre-trained models and employing transfer learning, companies can significantly cut down on the resources needed for training new models from scratch.

As awareness of environmental issues grows, so does the regulatory landscape. Companies must ensure that their ML practices comply with emerging regulations aimed at reducing environmental impact. This includes adhering to standards for energy efficiency, reporting carbon emissions, and implementing sustainable practices in their ML workflows.

Beyond individual companies, the broader ML community has a role to play in promoting environmentally informed practices. Open-source initiatives that develop and share energy-efficient algorithms and frameworks can help disseminate best practices. Collaborative efforts to create benchmarks for energy and resource efficiency in ML can drive industry-wide improvements.

Looking ahead, neuromorphic computing presents a promising avenue for achieving substantial energy savings in ML. Inspired by the architecture of the human brain, neuromorphic systems are designed to process information in a highly parallel and efficient manner [36–38]. Although still in the experimental stages, neuromorphic computing has the potential to revolutionize how we approach ML by drastically reducing power consumption and enhancing computational efficiency. Integrating such advanced technologies will further align ML practices with environmental sustainability goals.

In conclusion, environmentally informed ML is about integrating sustainability into the core of ML practices. By optimizing resource utilization, minimizing energy consumption, reducing the carbon footprint, and adhering to regulatory standards, companies can ensure that their ML models contribute positively to both economic and environmental goals. This holistic approach not only supports sustainable development, but also positions organizations as responsible and forward-thinking leaders in their fields.

5 Physically Informed

Physically informed ML integrates domain-specific knowledge into data-driven models, enhancing their performance and plausibility [12–14, 39]. This integration can be implemented at various stages of the ML pipeline, primarily categorized into pre-, in-, and post-processing methods. Pre-processing lays the foundations by acting on the data, in-processing embeds the knowledge by modifying the learning mechanisms, and post-processing aligns the outputs with domain expectations of the ML models. Each method plays a crucial role in ensuring the ML system is not only fed with high-quality data but also aligns its learning process and outcomes with domain-specific insights, thereby mitigating the "garbage in, garbage out" principle.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Pre-processing involves preparing and transforming the data before it is fed into an ML model. It addresses the quality of input data, ensuring that the ML model has the best possible starting point. Techniques such as data cleaning, feature engineering, and data augmentation fall under this category, where domain knowledge enhances the dataset's relevance and quality. Pre-processing capitalizes on domain expertise to navigate the ML model through the complex data landscape, minimizing the distance it needs to cover to generate valuable insights.

In-processing involves the direct incorporation of domain knowledge into the ML model's learning process. It requires deep integration of mathematical representations of domain insights, such as laws, trends, or constraints, into the learning algorithm itself. This could involve altering the model's functional form, introducing specific constraints, or embedding regularizers to maintain desirable properties like convexity and differentiability. The objective is to steer the model's learning mechanism in a way that benefits from domain knowledge and enhances predictive accuracy on a granular level. In-processing signifies a sophisticated blend of mathematical modeling and domain expertise to tune the learning process towards domain-aligned insights.

Post-processing refines the ML model's outputs to ensure they align with domain knowledge and expectations. It does not modify the ML model itself but adjusts its outputs through additional rules or models to enforce domain consistency. Techniques include using ML predictions as inputs to physical models for more controlled outcomes or applying logical rules to rectify inconsistencies in predictions. Post-processing leverages the existing ML capabilities and employs domain knowledge to contextualize and correct the model's predictions. This approach aims to mitigate potential errors and align the model's outputs with domain-specific truths, requiring substantial domain understanding to implement effectively.

Physically informed ML is a multifaceted field where various approaches enhance ML models by integrating domain knowledge at different stages of the ML pipeline. Through pre-processing, in-processing, and post-processing, domain expertise is systematically infused into the model's data, learning process, and outputs, respectively, improviing the model's predictive performances and its generalization bounds [40]. This integration not only improves model performance and accuracy but also ensures that the outcomes are meaningful and aligned with domain-specific realities. It highlights the importance of a synergistic collaboration between domain experts and data scientists, underpinning the successful application of informed ML methods.

6 Ethically Informed

Society at large requires ML models to operate within ethical boundaries, ensuring fairness, transparency, privacy, robustness, security, safety, and accountability [15–21]. This is crucial for maintaining public trust and addressing concerns related to bias, privacy, and the societal impact of ML-based technologies.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Ensuring fairness in ML involves addressing biases that may arise from training data or the modeling process. Bias can lead to discriminatory outcomes that disproportionately affect certain groups. Fair ML practices aim to eliminate these biases by implementing techniques such as algorithmic fairness, which includes methods like re-sampling, re-weighting, and adversarial debiasing. Regular audits and bias detection mechanisms are essential to identify and mitigate unfairness in ML models.

Transparency in ML models refers to the clarity and openness with which the workings of the model are communicated. This includes making the data sources, model architecture, and decision-making processes accessible and understandable to stakeholders. Techniques such as Explainable ML and model interpretability methods, including Shapley Additive Explanations and Local Interpretable Model-agnostic Explanations, play a crucial role in enhancing transparency. Transparent practices foster trust and allow for better scrutiny and understanding of ML systems. An interesting research line studies how to generalize explainability methods or define explainable-by-design models for complex data (see Section 2) [41, 42]. ML models often handle vast amounts of personal and sensitive data. Ensuring privacy involves safeguarding this data from unauthorized access and breaches. Techniques such as differential privacy, federated learning, and data anonymization are employed to protect individual privacy while still allowing for the effective training of ML models. Strict compliance with data protection regulations, such as the General Data Protection Regulation is also paramount in maintaining user privacy.

Robustness in ML models refers to their ability to maintain performance and functionality under various conditions, including adversarial attacks and noisy data. Developing robust models involves stress-testing them against potential threats and employing defensive techniques such as adversarial training and robustness regularization. Robust models ensure reliability and resilience, crucial for their deployment in real-world applications.

The security of ML models involves protecting them from adversarial attacks, data poisoning, and other threats that could compromise their integrity and functionality. Implementing secure coding practices, conducting regular security audits, and employing techniques such as adversarial detection and response are essential to safeguard ML systems. Ensuring security helps in maintaining the trustworthiness of ML models and protecting them from malicious exploitation.

Ensuring the safety of ML models is crucial, particularly when they are deployed in critical applications such as healthcare, autonomous driving, and finance. Safety involves a thorough testing and validation of the models to prevent unintended consequences and failures. Implementing fail-safe mechanisms, conducting rigorous scenario analysis, and adhering to industry safety standards are key practices to ensure the safe operation of ML systems.

Accountability in ML involves establishing mechanisms to attribute responsibility for the outcomes produced by ML models. This includes setting up governance frameworks that define the roles and responsibilities of various stakeholders involved in the ML lifecycle. Implementing audit trails, performance mon-

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

itoring, and accountability frameworks ensures that ML models operate within ethical boundaries and that any deviations can be promptly addressed. Accountability mechanisms are vital for fostering public trust and ensuring that ML systems are aligned with societal values.

In conclusion, ethically informed ML is essential for the responsible development and deployment of ML-based technologies. By addressing fairness, transparency, privacy, robustness, security, safety, and accountability, we can build ML models that not only perform effectively but also uphold ethical standards, ensuring their positive impact on society.

7 Conclusions

The exploration of ML in the context of complex data has revealed significant insights and potential advancements in the field. This special session highlights the critical need to integrate domain-specific knowledge into ML models to improve their generalization, sustainability, trustworthiness, reliability, security, and safety. By addressing the gap between purely data-driven models and the specific requirements and expertise of various domains, we can create more robust and effective ML solutions. Our special session underscores the importance of incorporating diverse forms of additional knowledge. For software developers, this means ensuring that ML models meet technical standards. For companies, it involves aligning ML practices with economic and environmental sustainability goals. Domain experts seek ML models that adhere to physical and logical laws, while society demands adherence to ethical principles. The contributions presented in this session illustrate the promising advances in Informed ML, where ML and interdisciplinary approaches converge to tackle real-world challenges. The showcased work demonstrates not only the potential of innovative ideas and improvements but also the limitations that must be addressed to achieve practical and impactful solutions. In conclusion, bridging the gap between data-driven approaches and domain-specific knowledge is essential for the evolution of ML. This integration will pave the way for more reliable, secure, and ethically sound applications, ultimately contributing to the progress of society in the contemporary era of data-driven decision-making. The early findings and contributions discussed here lay a strong foundation for future research and development in Informed ML for Complex Data.

References

[1] J. Jumper, R. Evans, A. Pritzel, T. Green, and Others. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

[2] OpenAI. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

[3] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, and Others. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 2017.

[4] L. Oneto, N. Navarin, B. Biggio, F. Errica, and Others. Towards learning trustworthily, automatically, and with guarantees on graphs: An overview. *Neurocomputing*, 493:217–243, 2022.

[5] C. Zhou, Q. Li, C. Li, J. Yu, and Others. A comprehensive survey on pretrained foundation models: A history from bert to chatgpt. *arXiv preprint arXiv:2302.09419*, 2023.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

[6] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 32(1):4–24, 2020.

[7] D. Bacciu, F. Errica, A. Micheli, and M. Podda. A gentle introduction to deep learning for graphs. *Neural Networks*, 129:203–221, 2020.

[8] E. Giunchiglia, F. Imrie, M. Van Der Schaar, and T. Lukasiewicz. Machine learning with requirements: A manifesto. *Neurosymbolic Artificial Intelligence*, (in press), 2024.

[9] A. Van Wynsberghe. Sustainable ai: Ai for sustainability and the sustainability of ai. *AI and Ethics*, 1(3):213–218, 2021.

[10] C. Isensee, K. M. Griese, and F. Teuteberg. Sustainable artificial intelligence: A corporate culture perspective. *Sustainability Management Forum— NachhaltigkeitsManagement-Forum*, 29(3):217–230, 2021.

[11] J. Khakurel, B. Penzenstadler, J. Porras, A. Knutas, and . Zhang. The rise of artificial intelligence under the lens of sustainability. *Technologies*, 6(4):100, 2018.

[12] L. Von Rueden, S. Mayer, K. Beckh, B. Georgiev, and Others. Informed machine learning–a taxonomy and survey of integrating prior knowledge into learning systems. *IEEE Transactions on Knowledge and Data Engineering*, 35(1):614–633, 2021.

[13] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.

[14] F. Leiser, S. Rank, M. Schmidt-Kraepelin, S. Thiebes, and A. Sunyaev. Medical informed machine learning: A scoping review and future research directions. *Artificial Intelligence in Medicine*, 145:102676, 2023.

[15] A. F. Winfield, K. Michael, J. Pitt, and V. Evers. Machine ethics: The design and governance of ethical ai and autonomous systems [scanning the issue]. *Proceedings of the IEEE*, 107(3):509–517, 2019.

[16] L. Oneto, S. Ridella, and D. Anguita. Towards algorithms and models that we can trust: a theoretical perspective. *Neurocomputing*, (in press), 2024.

[17] B. Li, P. Qi, B. Liu, S. Di, and Others. Trustworthy ai: From principles to practices. *ACM Computing Surveys*, 55(9):1–46, 2023.

[18] B. Biggio and F. Roli. Wild patterns: Ten years after the rise of adversarial machine learning. *Pattern Recognition*, 84:317–331, 2018.

[19] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan. A survey on bias and fairness in machine learning. *ACM computing surveys*, 54:1–35, 2021.

[20] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, and Others. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information fusion*, 58:82–115, 2020.

[21] M. Al-Rubaie and J. M. Chang. Privacy-preserving machine learning: Threats and solutions. *IEEE Security & Privacy*, 17(2):49–58, 2019.

[22] C. Ying, T. Cai, S. Luo, S. Zheng, G. Ke, D. He, Y. Shen, and T.Y. Liu. Do transformers really perform badly for graph representation? In A. Beygelzimer, Y. Dauphin, P. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, 2021.

[23] F. Petruzzellis, A. Testolin, and A. Sperduti. Assessing the emergent symbolic reasoning abilities of llama large language models. In *ICANN 2024*. arXiv:2406.06588 [cs].

[24] Z. Wu, S. Pan, F. Chen, G. Long, and C. Zhang et al. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1), 2021.

[25] D. Bacciu, F. Errica, A. Micheli, and M. Podda. A gentle introduction to deep learning for graphs. *Neural Networks*, 129, 2020.

[26] M. Li, A. Micheli, Y. G. Wang, S. Pan, P. Lió, G. S. Gnecco, and M. Sanguineti. Guest editorial: Deep neural networks for graphs: Theory, models, algorithms, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 35(4):4367–4372, 2024.

[27] N. Navarin, L. Pasa, and A. Sperduti. Towards the application of backpropagation-free graph convolutional networks on huge datasets. In *ESANN 2024 proceesdings*, Bruges (Belgium) and online, 2024. Ciaco - i6doc.com.

[28] D. Tortorella and A. Micheli. Continual learning with graph reservoirs: Preliminary experiments in graph classification. In *ESANN 2024 proceesdings*, Bruges (Belgium) and

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

online, 2024. Ciaco - i6doc.com.

[29] M. M. Bronstein, J. Bruna, Y. Lecun, A. Szlam, and P. Vandergheynst. Geometric deep learning: Going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

[30] D. Bacciu and F. Landolfi. Generalizing convolution to point clouds. In *ESANN 2024 proceesdings*, Bruges (Belgium) and online, 2024. Ciaco - i6doc.com.

[31] D. Sculley, G. Holt, D. Golovin, E. Davydov, and Others. Hidden technical debt in machine learning systems. *Neural information processing systems*, 2015.

[32] E. Strubell, A. Ganesh, and A. McCallum. Energy and policy considerations for modern deep learning research. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 13693–13696, 2020.

[33] K. Nakajima and I. Fischer. *Reservoir computing*. Springer, 2021.

[34] M. Yan, C. Huang, P. Bienstman, P. Tino, W. Lin, and J. Sun. Emerging opportunities and challenges for the future of reservoir computing. *Nature Communications*, 15(1):2056, 2024.

[35] C. Gallicchio. Euler state networks: Non-dissipative reservoir computing. *Neurocomputing*, 579:127411, 2024.

[36] C. D. Schuman, T. E. Potok, R. M. Patton, J. D. Birdwell, M. E. Dean, G. S. Rose, and J. S. Plank. A survey of neuromorphic computing and neural networks in hardware. *arXiv preprint arXiv:1705.06963*, 2017.

[37] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier. Physics for neuromorphic computing. *Nature Reviews Physics*, 2(9):499–510, 2020.

[38] G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose. Recent advances in physical reservoir computing: A review. *Neural Networks*, 115:100–123, 2019.

[39] N. Navarin, P. Frazzetto, L. Pasa, P. Verzelli, F. Visentin, A. Sperduti, and C. Alippi. Physics-informed graph neural cellular automata: an application to compartmental modelling. In *2024 International Joint Conference on Neural Networks (IJCNN)*, 2024.

[40] L. Oneto, D. Anguita, and S. Ridella. Informed machine learning: Excess risk and generalization. In *ESANN 2024 proceesdings*, Bruges (Belgium) and online, 2024. Ciaco - i6doc.com.

[41] M. Fontanesi, A. Micheli, and M. Podda. Xai and bias of deep graph networks. In *ESANN 2024 proceesdings*, Bruges (Belgium) and online, 2024. Ciaco - i6doc.com.

[42] F. Spinnato, A. Cossu, R. Guidotti, A. Ceni, C. Gallicchio, and D. Bacciu. Enhancing echo state networks with gradient-based explainability methods. In *ESANN 2024 proceesdings*, Bruges (Belgium) and online, 2024. Ciaco - i6doc.com.