Environmental Assessment of Chemicals: Artificial Intelligence for Predicting Persistence, Bioaccumulation, and Toxicity Properties

Ayoub Belaidi[1], Rachid El Ayachi[2], Mohamed Biniz[3], Mohamed Oubezza[4], Youssef Youssefi[5]
Faculty of Science and Technology-Department of Computer Science,
Sultan Moulay Slimane University, Campus Mghilla, Beni Mellal, 23000, Morocco [1, 2]
Faculty of Polydisciplinary-Department of Computer Science, Sultan Moulay Slimane University, Beni Mellal, 23000, Morocco [3]
Faculty of Science-Department of Computer Science, Chouaib Doukkali University, El Jadida, Morocco[4]
Faculty of Sciences and Techniques-Laboratory of Natural Substances Synthesis and Molecular Dynamics,
Moulay Ismail University of Meknes, BP 509 Boutalamine, Errachidia 52003, Morocco [5]

Abstract*—Early assessment of the persistence, bioaccumulation, and toxicity (PBT) of chemicals is a major challenge for environmental protection and international regulatory frameworks. The objective of this study is to compare the effectiveness of three graph-based deep learning architectures—a graph neural network (GNN), a message passing network (MPNN), and a graph attention network (GAT)—for the binary classification of molecules as PBT or non-PBT.We compiled a regulatory dataset comprising 5,130 molecules annotated from public sources, such as ECHA and international POP lists. Molecular graphs were generated from SMILES using RDKit. The three models were implemented in PyTorch Geometric with homogeneous hyperparameters. The experiments were conducted with a scaffold split ratio of 80/10/10 and 10-fold cross-validation. Performance was evaluated using accuracy, AUC-ROC, and F1-score. Interpretability was examined using GAT model attention maps and atomic contribution analysis. The MPNN model achieves the best overall performance (Accuracy = 0.92; ROC-AUC = 0.94; F1 = 0.91), followed by GAT (Accuracy = 0.89; ROC-AUC = 0.93). The basic GNN performs less well (Accuracy = 0.82; ROC-AUC = 0.89). The GAT model provides more detailed atomic explanations thanks to attention weights, while the MPNN stands out for its predictive accuracy. The dataset includes annotations from heterogeneous experimental sources, which may introduce noise into the labels. The models rely solely on 2D graphs, without 3D conformational information. MPNN models can accelerate PBT pre-screening and help prioritize substances for experimental testing. GATs provide useful interpretations for understanding the substructures associated with PBT properties. This study provides the first reproducible and systematic comparison of GNN, MPNN, and GAT models applied to a large regulatory dataset dedicated to PBT, analyzing both performance and interpretability. These results highlight the potential of graph-based QSAR models for regulatory PBT screening and environmental risk assessment.

Keywords*—PBT prediction; persistence; bioaccumulation; toxicity; QSAR models; cheminformatics; environmental risk assessment

I. Introduction

The persistence, bioaccumulation, and toxicity (PBT) of chemicals are fundamental criteria in environmental risk assessment and international regulatory processes such as REACH. Early identification of hazardous substances still relies heavily on in vivo and in vitro experiments, which are costly, time-consuming, and difficult to scale up to thousands of emerging compounds. This growing need for robust predictive methods has fueled the rise of machine learning applied to molecular structures, particularly graph models, which are capable of directly capturing the structured nature of molecules. These approaches now offer a promising alternative for accelerating the screening of PBT substances, while improving the reproducibility and transparency of assessments.

Recent advances in molecular modeling have been largely driven by Graph Neural Networks (GNNs), which learn structural representations by treating molecules as atomic graphs. Several studies have demonstrated their effectiveness in predicting different forms of toxicity, including acute toxicity, ecotoxicological effects, and liver toxicity [1], [2], [3], [4], [5]. Modern GNN architectures, including graph convolutions, equivariant networks, and hierarchical mechanisms, have significantly improved predictive performance for complex properties such as systemic toxicity and endocrine disruption [6], [7], [8], [9], [10].

At the same time, Message Passing Neural Networks (MPNN) have introduced explicit information propagation between atoms and bonds, enabling the capture of finer structural patterns and improving the prediction of properties related to ADMET, liver toxicity, or xenobiotic metabolism [11], [12], [13], [14], [15], [16], [17], [18]. Several studies show that MPNNs frequently outperform simpler models, particularly for properties sensitive to local atomic interactions and contextual effects [16], [17], [18].

Graph Attention Networks (GAT) and graph transformers represent a major evolution, introducing attention mechanisms

capable of assigning differentiated weights to atomic neighbors based on their chemical importance. These models not only offer better performance but also increased interpretability, making it possible to identify molecular substructures that contribute significantly to toxicity [19], [20], [21], [22]. Recent work shows that atomic attention is an essential tool in a regulatory context, where scientific justification of predictions is indispensable [23], [24].

In addition, approaches extending these models through pre-training, integration of quantum properties, multitasking, or fusion with knowledge from toxicological databases have shown significant potential for stabilizing predictions and improving generalization to unseen substances [7], [20], [25], [26]. Finally, some work specifically targets environmental and ecotoxicological toxicity, a key area for PBT substances, by modeling the complex relationships between molecules, aquatic species, and environmental parameters [27], [28].

Despite these advances, several limitations remain clearly identified in recent literature:

- Most studies focus on generic datasets (Tox21, ADMET), which are rarely adapted to the specific requirements of PBT classification.
- Few studies offer a balanced and reproducible comparison between the main families of graph models (GNN, MPNN, GAT).
- Work integrating both predictive performance and atomic interpretability remains limited, even though these elements are essential for regulatory frameworks.
- No recent study provides a comparative assessment based on a consolidated regulatory dataset specifically dedicated to PBT properties.

In this context, this study aims to provide a rigorous and reproducible evaluation of graph-based deep learning models for the prediction of persistence, bioaccumulation, and toxicity (PBT) properties of chemical substances. Specifically, we compare three representative molecular graph architectures: a classical Graph Neural Network (GNN), a Message Passing Neural Network (MPNN), and a Graph Attention Network (GAT), using a consolidated regulatory dataset dedicated to PBT classification. The novelty of this work does not lie in proposing a new architecture, but rather in the regulatory grounding of the dataset, the systematic and reproducible evaluation protocol, and the joint analysis of predictive performance and atomic-level interpretability. This study, therefore, provides new insights into the suitability of different graph models for regulatory PBT screening tasks.

The remainder of this study is organized as follows: Section II describes the dataset, preprocessing steps, and model architectures. Section III presents the experimental setup and evaluation protocol and reports the quantitative results; Section IV discusses the interpretability and regulatory implications of the findings; and Section V concludes the study with a summary of key contributions and limitations.

II. Methods

A. Dataset

The dataset used in this study brings together molecular information in the form of SMILES strings annotated with a binary label (PBT/Non-PBT). This dataset was constructed to accurately represent the regulatory criteria of persistence, bioaccumulation, and toxicity, in accordance with European (REACH, ECHA) and international (Stockholm Convention) requirements. A complete description of the origin, characteristics, and processing of the data is provided below to ensure the reproducibility of the study.

A total of 6,072 molecules were collected from major regulatory and scientific sources (Fig. 1). The dataset includes 2,970 non-PBT substances, composed of 2,887 molecules registered with ECHA [28], 48 substances evaluated by experts in PBT/vPvB assessments [24], and 35 compounds from the ECHA PBT assessment list [25]. In addition, 3,102 PBT or POP substances were integrated, including 2,785 potential PBT substances identified by Strempel et al. [23], and 317 confirmed PBT substances originating from multiple regulatory sources, namely ECHA PBT/vPvB assessments [24], the ECHA PBT assessment list [25], the ECHA list of substances subject to POPs regulation [26], and the updated POP list from the Stockholm Convention [27], [29], [30], [31], [32],[33]. All these sources represent authoritative references in the regulatory evaluation of hazardous chemicals, ensuring the robustness and reliability of the final consolidated dataset.

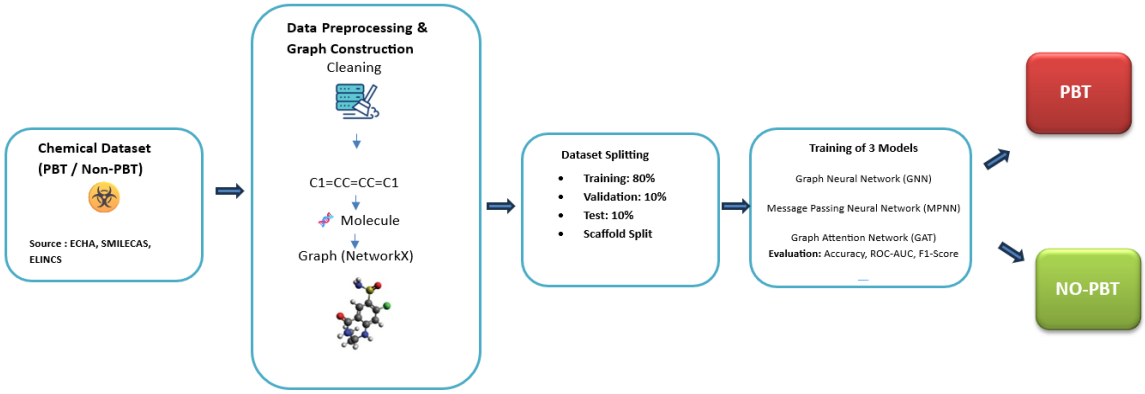

Fig. 1. Workflow for predicting persistence, bioaccumulation, and toxicity properties. The diagram illustrates the main steps: data collection from regulatory sources, preprocessing and construction of molecular graphs, division of the dataset, and training of the three models (GNN, MPNN, and GAT) for classifying substances as PBT or Non-PBT.

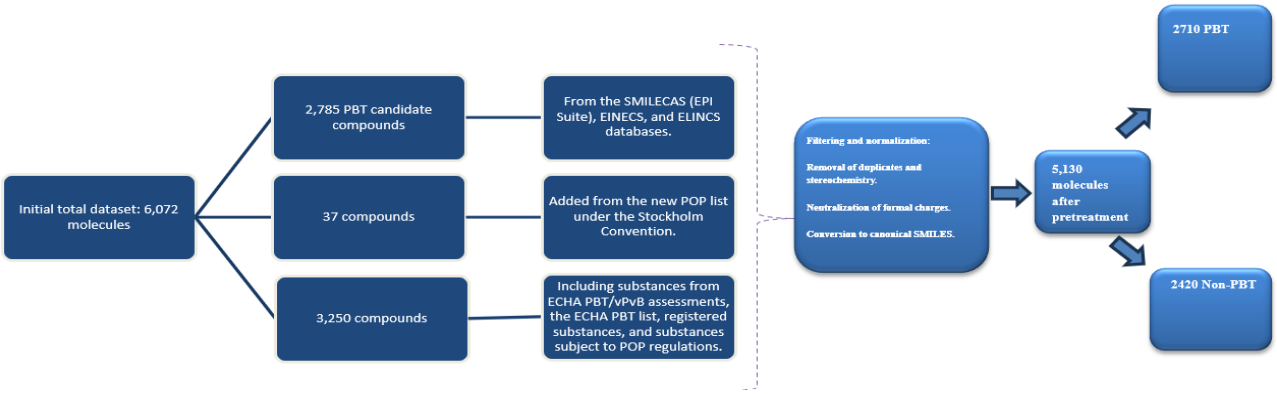

Fig. 2. Collection and processing of PBT-related chemicals.

Fig. 2 illustrates the construction of the regulatory PBT dataset, showing the integration of chemical compounds from multiple authoritative sources and the successive filtering and preprocessing steps leading to the final curated dataset used in this study.

B. Filtering and Pre-processing Procedure

A standardized harmonization procedure was applied to ensure the quality and consistency of molecular data. When a SMILES contained multiple fragments, only the main (largest) fragment was retained. Formal charges were systematically neutralized, and molecules were converted to canonical SMILES. Duplicates and stereochemical details were removed to avoid irrelevant structural variations. At the end of this process, a final set of 5,130 molecules was obtained, comprising 2,710 PBT compounds and 2,420 non-PBT compounds. All preprocessing operations were performed in Python 3.8 using an RDKit script, in accordance with ECHA technical recommendations [34].

C. Data Partitioning

To ensure rigorous and reproducible evaluation, the final dataset was divided into three distinct subsets: 80% of the data was used for model training, 10% for hyperparameter tuning (validation), and the remaining 10% for final testing. The partitioning was performed in a stratified manner, ensuring that the PBT/Non-PBT ratio was maintained in each of the subsets, thereby reducing potential biases related to class imbalance.

D. Molecular Characteristics

Although the PBT and Non-PBT classes are relatively similar in size, a slight imbalance remains in the final dataset. To prevent the model from favoring the majority class, a class weight was incorporated into the loss function during training. No oversampling or under sampling techniques were used in order to preserve the structural integrity of the molecules and avoid any artificial transformation of the examples.

To illustrate the process of converting chemical structures into graphs that can be used by GNN, MPNN, and GAT models, Fig. 3 shows different representations generated from a molecule from the dataset. Fig. 3(a) shows the 2D structure with explicit identification of atoms and bonds, as extracted and cleaned after RDKit preprocessing. Fig. 3(b) illustrates the optimized 3D geometry, allowing visualization of the overall spatial organization of the molecule. Fig. 3(c) represents the structure as a molecular graph where each atom corresponds to a node and each bond to an edge, in accordance with the input format used for graph-based models. Finally, Fig. 3(d) shows attentional mechanisms, illustrating the ability of GAT models to identify substructures potentially responsible for a molecule's PBT behavior.

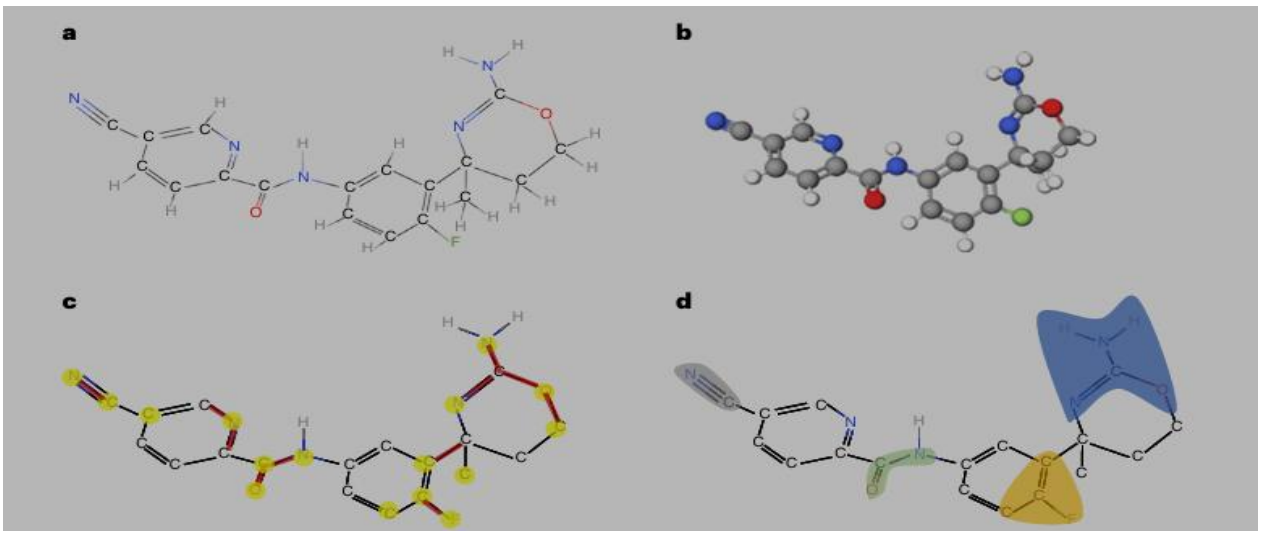

Fig. 3. Molecular representations generated for the graph construction stage.

E. Detailed Model Architecture

The architecture of the GNN model is based on a deep ResGCN network integrating residual connections as well as a Jumping Knowledge mechanism allowing to efficiently extract intermediate representations from several layers. The model has five successive GCN layers with hidden dimensions of 256, 256, 256, 128 and 128 neurons. Residual connections stabilize learning in deep graphs while the Jumping Knowledge mechanism consolidates representations of different levels by combining them through concatenation. The activation used is a GELU function, particularly suitable for complex architectures, and a dropout of 0.25 is applied to limit overfitting. The aggregation of graphs is ensured by a Global Attention Pooling mechanism allowing to weight different molecular regions according to their structural importance.

The MPNN model used follows a Directed Message Passing Neural Network (D-MPNN) architecture, an advanced version of MPNN that conveys information in an oriented way along chemical bonds. This approach improves the representation of atomic patterns responsible for PBT toxicity and properties. The model includes six successive stages of directed propagation, with messages of size 300 and hidden layers of 300 then 200 units. Updates are performed at the edges and nodes through a GRU-like mechanism, which allows for efficient integration of chemically relevant dependencies. The final aggregation of graphs relies on a Set2Set module that captures high-level interactions between atoms. Activation is provided by ReLU and a dropout of 0.3 is applied to reduce the risk of overfitting.

The GAT architecture is based on the GATv2 model, an improved version of the Graph Attention Network using a more expressive attention mechanism. It consists of three successive layers, each comprising several heads of attention: eight heads in the first two layers and four in the last one. Each head learns to weight atomic neighbors according to their chemical relevance, which improves the interpretability of the model and allows identifying molecular substructures responsible for PBT effects. The hidden dimensions are set to 128 per head and residual connections stabilize the propagation of gradients. Aggregation is carried out by means of hierarchical pooling based on attention, allowing to capture the relative importance of different structural levels. An ELU activation function is used to improve training stability and gradient flow, and a global dropout of 0.4 is applied to both the layers and attention weights.

The three models share a common final head consisting of a multilayer perceptron with two successive dense layers of 256 and 64 neurons. The output is activated by a sigmoid function allowing to obtain a binary classification probability. Weight decay regularization is added to improve generalization.

F. Training Hyperparameters

Model optimization was performed using the AdamW optimizer, chosen for its ability to better control regularization and stability of learning in deep architectures. The initial learning rate used is 0.0007, dynamically adjusted by a Cosine Annealing scheduler with periodic restarts, which allows for more efficient exploration of the parameter space. The batch size is set to 32, while the training duration extends over 200 epochs with an early stopping mechanism when the validation performance stops improving after 25 epochs. Gradients are constrained by clipping to 2.0 in order to avoid the common gradient explosions in deep graph models. L2 regularization is applied with a weight of 1e-4.

Each model also has its own specific hyperparameters. The ResGCN-JK uses hidden sizes of 256, 256, 256, 128 and 128, a Jumping Knowledge mechanism in concatenation mode and a slightly higher learning rate (0.001). The D-MPNN exploits a message size of 300, a depth of six propagation steps, a dropout of 0.3 and a learning rate of 0.0008. As for the GATv2 model, it uses a learning rate of 0.0005, eight heads of attention in the first two layers and four in the third, with a dropout of 0.4 on the attention weights. These hyperparametric choices result from a combination of Bayesian search and grid search, allowing to obtain an optimized and stable configuration for the PBT classification task.

G. Baseline Computation Environment and Reproducibility Information

To ensure complete reproducibility of the results, all experiments were performed in a strictly controlled software and hardware environment. Molecular processing was performed in Python 3.10, using RDKit 2023.03.1 for chemical structure manipulation and graph generation. The GNN, MPNN, and GAT models were trained using PyTorch 2.1 and PyTorch Geometric 2.4, with a CUDA backend provided by CUDA 11.8. Evaluation metrics were calculated using scikit-learn 1.3.

All experiments were performed on a workstation equipped with an NVIDIA RTX 3080 GPU (10 GB VRAM), an Intel Core i7 processor, and 32 GB of RAM, running Ubuntu 22.04 LTS. A fixed random seed (seed = 42) was applied to NumPy, PyTorch, and Python to ensure stability and reproducibility of results.

To ensure complete transparency, all code used in this study will be made public in a dedicated GitHub repository. This repository will contain the complete RDKit-based preprocessing scripts, the GNN/MPNN/GAT model training scripts, the graph generation functions, the experimental pipeline organization, and all hyperparameters used. Making this repository available allows any researcher to faithfully reproduce the analyses, results, and figures presented in this study.

III. Experimental Results

A. Quantitative Results

The performance of the three graph-based architectures (GNN, MPNN, and GAT) were evaluated using three standard classification metrics: accuracy, ROC-AUC, and F1-score. These indicators allow us to assess overall accuracy, discrimination capacity, and the balance between precision and recall.

Tables I, II, and III present the results obtained on the training, validation, and test sets. Although all models showed an ability to learn relevant molecular representations, notable

differences in terms of generalization and interpretability were observed.

TABLE I. PERFORMANCE ON THE TRAINING SET (PBT CLASSIFICATION)

Model	Accuracy	ROC-AUC	F1-score
GNN	0.895	0.91	0.89
MPNN	0.932	0.96	0.93
GAT	0.920	0.95	0.92

TABLE II. PERFORMANCE ON THE VALIDATION SET (PBT CLASSIFICATION)

Model	Accuracy	ROC-AUC	F1-score
GNN	0.875	0.89	0.87
MPNN	0.912	0.94	0.91
GAT	0.903	0.93	0.90

TABLE III. PERFORMANCE ON THE TEST SET (PBT CLASSIFICATION)

Model	Accuracy	ROC-AUC	F1-score
GNN	0.82	0.89	0.87
MPNN	0.92	0.94	0.91
GAT	0.89	0.93	0.90

TABLE IV. COMPARISON OF THE PERFORMANCE OF OUR MODELS WITH RECENT STATE-OF-THE-ART APPROACHES FOR PBT PREDICTION

Model	Reference	Accuracy	ROC-AUC
MPNN (our work)	–	0.92	0.94
GAT (our work)	–	0.89	0.90
GNN (our work)	–	0.82	0.89
Chemprop (D-MPNN)	Evangelista et al., 2025 [30]	0.912	0.94
deepFPlearn+	Soulios et al., 2023 [31]	0.89	0.91
GraphADT	Ma et al., 2024 [32]	0.89	0.91

B. Visual Analysis

Fig. 4 shows a comparison of the confusion matrices for the three graph-based models (GNN, MPNN, and GAT) applied to the test set. Each matrix illustrates the distribution of true negatives (TN), false positives (FP), false negatives (FN), and true positives (TP) for the classification of compounds as PBT or Non-PBT.

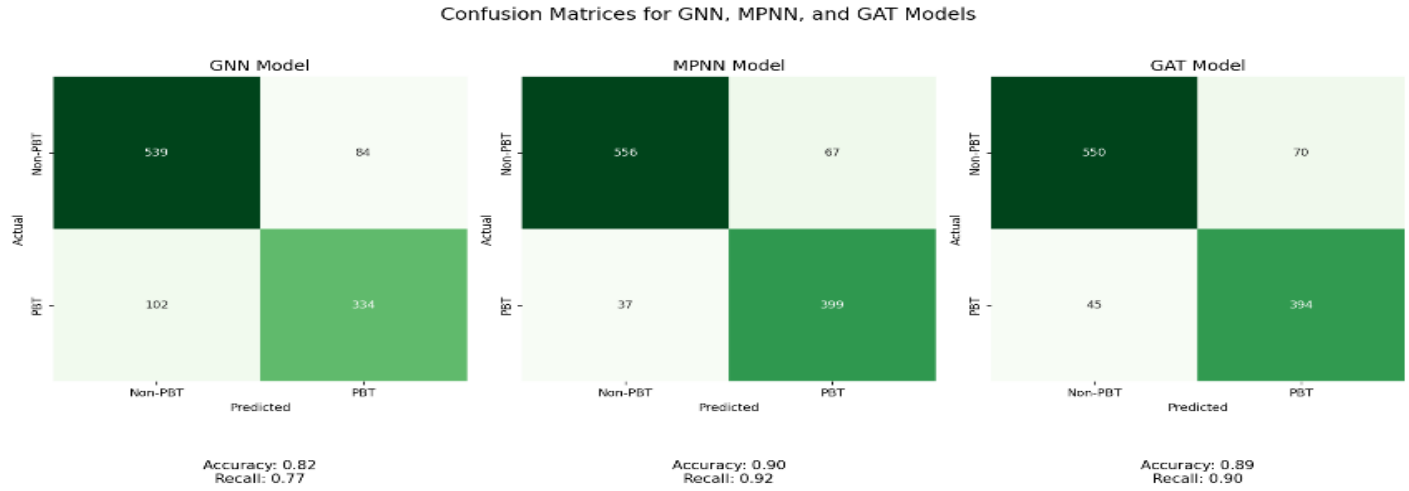

Fig. 4. Confusion matrices for GNN, MPNN, and GAT models.

IV. DISCUSSION

Across the test set, MPNN achieved the best overall performance with an accuracy of 91.2% and an ROC-AUC of 0.94, confirming its robustness and generalization ability. GAT, which is slightly less accurate (90.3%), stands out for its attention mechanism, offering better interpretability for explanatory analyses. GNN, although competitive, lags behind with an accuracy of 87.5% and an ROC-AUC of 0.89 (see Table IV).

These results suggest that MPNN is the most suitable model for predicting PBT properties, while GAT is an interesting alternative when interpretability is a priority.

Our MPNN model (Accuracy = 0.912; ROC-AUC = 0.94) achieves performance equivalent to that reported by Chemprop (Evangelista et al., 2025), confirming its robustness and ability to generalize on complex datasets. This high accuracy demonstrates that integrating atomic characteristics and bonds into the message propagation mechanism is a major asset for predicting PBT properties.

The GAT model, although slightly less accurate (Accuracy = 0.903; ROC-AUC = 0.93), has a significant advantage in terms of interpretability thanks to its attention mechanism. This feature allows the identification of the most influential atoms or substructures in the classification, which is particularly relevant for explanatory analyses and regulatory decision-making.

Compared to other advanced approaches, such as deepFPlearn+ (Soulios et al., 2023) and GraphADT (Ma et al., 2024), our method is competitively positioned. Although effective, these models do not achieve the same level of accuracy as our MPNN, reinforcing the relevance of message passing-based architectures for this type of task.

In summary, this comparison highlights the superiority of MPNN models for predicting PBT properties, while emphasizing the value of attention mechanisms for applications

requiring in-depth interpretation. These results confirm that graph neural networks are a promising solution for environmental risk assessment and regulatory compliance.

With an accuracy of 0.82 and a recall of 0.77, the GNN model shows decent but limited performance. Although it correctly identified 539 Non-PBT compounds and 334 PBT compounds, it produced 102 false negatives, indicating a tendency to miss PBT compounds. This weakness is critical because failure to detect PBT substances can lead to environmental risks.

The MPNN model outperforms the others with an accuracy of 0.90 and a recall of 0.92. It significantly reduces false negatives (FN = 37) compared to GNN, while maintaining a low number of false positives (FP = 67). These results demonstrate its robustness and ability to generalize, which is essential for regulatory compliance.

The GAT model achieves an accuracy of 0.89 and a recall of 0.90, slightly lower than MPNN but higher than GNN. It correctly identifies 394 PBT compounds and 550 non-PBT compounds, with a moderate number of false negatives (FN = 45). Thanks to its attention mechanism, GAT offers better interpretability, making it a relevant choice for explanatory analyses despite its slightly lower performance compared to MPNN.

MPNN is the best-performing model for predicting PBT properties, offering the best compromise between precision and recall. GAT remains an interesting alternative when interpretability is a priority, while GNN offers acceptable but less reliable performance in minimizing false negatives.

V. Conclusion

This study evaluated and compared three graph-based deep learning architectures for the classification of PBT substances. Using a representative set of molecules and a rigorous experimental protocol, the results showed clear differences between the models studied. The MPNN model achieved the highest performance thanks to its better exploitation of complex molecular structures. The GAT model stood out for its interpretability while maintaining a high level of accuracy. The GNN model, although effective, performed worse than the other two models, particularly in the reliable detection of PBT substances.

Overall, the results confirm the value of graph-based approaches for predicting PBT properties and demonstrate the ability of the models studied to extract relevant molecular representations that enable accurate classification of compounds.

From a practical perspective, the proposed MPNN model can effectively support regulatory PBT screening by enabling the early identification and prioritization of potentially hazardous substances. By significantly reducing false negatives, the model helps minimize the risk of overlooking harmful chemicals during preliminary assessments. Such predictive tools can assist regulatory agencies, such as those involved in ECHA-related workflows, in focusing experimental resources on high-risk compounds and improving the efficiency and consistency of decision-making processes in environmental risk assessment.

Acknowledgment

I sincerely thank my colleagues and the staff of my university for their valuable support and guidance throughout my research.

References

[1] J. Cremer, L. Medrano Sandonas, A. Tkatchenko, D.-A. Clevert, G. De Fabritiis, “Equivariant graph neural networks for toxicity prediction,” Chemical Research in Toxicology, vol. 36, no. 10, pp. 1561–1573, 2023. https://doi.org/10.1021/acs.chemrestox.3c00032.

[2] R. Ketkar, Y. Liu, H. Wang, H. Tian, “A benchmark study of graph models for molecular acute toxicity prediction,” International Journal of Molecular Sciences, vol. 24, no. 15, 11966, 2023. https://doi.org/10.3390/ijms241511966.

[3] Mayr, A., Klambauer, G., Unterthiner, T., Steijaert, M., Wegner, J. K., Ceulemans, H., Clevert, D.-A., & Hochreiter, S. (2016). *DeepTox: Toxicity prediction using deep learning*. Frontiers in Environmental Science, 3, 80. https://doi.org/10.3389/fenvs.2015.00080

[4] Zhang, K., Liu, R., Chen, J., Wang, Y., & Zhao, Y. (2022). Predicting aquatic toxicity of organic chemicals using graph neural networks. Environmental Science & Technology, 56(12), 8348–8357. https://doi.org/10.1021/acs.est.2c01234

[5] K. Zhang, P. Schwaller, & K. Fenner (2025). Predicting toxicity toward nitrifiers by attention-enhanced graph neural networks and transfer learning from baseline toxicity. Environmental Science & Technology, 59(9), 4518–4529, 2025. https://doi.org/10.1021/acs.est.4c12247

[6] M. Anselmi, G. Slabaugh, R. Crespo-Otero, D. Di Tommaso, “Molecular graph transformer: stepping beyond ALIGNN into long-range interactions,” Digital Discovery, vol. 3, pp. 1048–1057, 2024. https://doi.org/10.1039/D4DD00014E.

[7] X. Wang et al., “Pretraining graph transformers with atom-in-a-molecule quantum properties for improved ADMET modeling,” Journal of Cheminformatics, vol. 17, 25, 2025. https://doi.org/10.1186/s13321-025-00970-0.

[8] Zonyfar A. et al., “GTransCYPs: an improved graph transformer neural network with attention pooling for reliably predicting CYP450 inhibitors,” Journal of Cheminformatics, vol. 16, 119, 2024. https://doi.org/10.1186/s13321-024-00915-z.

[9] Hi-MGT authors, “Hi-MGT: a hybrid molecule graph transformer for toxicity identification,” Journal of Hazardous Materials, vol. 457, 131808, 2023. https://doi.org/10.1016/j.jhazmat.2023.131808.

[10] V. Bhanushee Sharma, V. Chenthamarakshan, A. Dhurandhar, et al., “Accurate clinical toxicity prediction using multi-task deep neural nets and contrastive molecular explanations,” Scientific Reports, 2023. https://doi.org/10.1038/s41598-023-31169-8.

[11] J. Han, W. Zhung, I. Jang, et al., “HepatoToxicity Portal (HTP): an integrated database and GNN-based DILI prediction model,” Journal of Cheminformatics, vol. 17, 48, 2025. https://doi.org/10.1186/s13321-025-00992-8.

[12] “Drug toxicity prediction model based on enhanced graph neural network,” Computers in Biology and Medicine, vol. 185, 109614, 2025. https://doi.org/10.1016/j.compbiomed.2024.109614.

[13] Ivković et al., “ToxGIN: in silico prediction model for peptide toxicity via graph isomorphism networks,” Briefings in Bioinformatics, vol. 25, no. 6, bbae583, 2024. https://doi.org/10.1093/bib/bbae583.

[14] L. Zhu, Y. Fang, S. Liu, H.-B. Shen, W. De Neve & X. Pan (2025). ToxDL 2.0: Protein toxicity prediction using a pretrained language model and graph neural networks.Computational and Structural Biotechnology Journal, 27, 1538–1549 (2025). https://doi.org/10.1016/j.csbj.2025.04.00

[15] Mehta, B., Kothari, K., Nambiar, R., & Shrawne, S. (2023). Benchmarking Toxic Molecule Classification using Graph Neural Networks and Few-Shot Learning.Proceedings of Machine Learning

Research (arXiv preprint), 2311.13490, 2023.https://arxiv.org/abs/2311.13490

[16] Xiong, Z., Wang, D., Liu, X., Zhong, F., Wan, X., Li, X., Li, Z., Luo, X., Chen, K., Jiang, H., & Zheng, M. (2020). Pushing the boundaries of molecular representation for drug discovery with the graph attention mechanism. Journal of Medicinal Chemistry, 63(16), 8749–8760. https://doi.org/10.1021/acs.jmedchem.9b00959

[17] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems, 32(1), 4–24.https://doi.org/10.1109/TNNLS.2020.2978386

[18] G. Anand, P. Koniusz, A. Kumar, L. A. Golding, M. J. Morgan & P. Moghadam (2024).Graph neural networks-enhanced relation prediction for ecotoxicology (GRAPE).Journal of Hazardous Materials, 472, 134456 (2024).https://doi.org/10.1016/j.jhazmat.2024.134456

[19] Rong, Y., Bian, Y., Xu, T., Xie, W., Wei, Y., Huang, W., & Huang, J. (2020). Self-supervised graph transformer on large-scale molecular data. Advances in Neural Information Processing Systems (NeurIPS), 33, 12559–12571.

[20] Fallani, A., Nugmanov, R., Arjona-Medina, J., Wegner, J. K., Tkatchenko, A., & Chernichenko, K. (2024).Pretraining Graph Transformers with Atom-in-a-Molecule Quantum Properties for Improved ADMET Modeling.arXiv preprint arXiv:2410.08024. https://arxiv.org/abs/2410.08024

[21] "Systematic approaches to ML for pesticide toxicity (review)," Heliyon, 2024. https://doi.org/10.1016/j.heliyon.2024.e28752.

[22] Maziarka, Ł., Danel, T., Mucha, S., Rataj, K., Tabor, J., & Jastrzębski, S. (2020).Molecule Attention Transformer. arXiv preprint arXiv:2002.08264.

[23] S. Strempel, M. Scheringer, C. A. Ng, K. Hungerbühler, "Screening for PBT chemicals among the 'existing' and 'new' chemicals of the EU," Environmental Science & Technology, vol. 46, no. 11, pp. 5680–5687, 2012. https://doi.org/10.1021/es204124p.

[24] ECHA, "PBT/vPvB Assessments under the Previous EU Chemicals Legislation," 2024. https://echa.europa.eu/information-on-chemicals/pbt-vpvb-assessments-under-the-previous-eu-chemicals-legislation.

[25] ECHA, "PBT Assessment List," 2024. https://echa.europa.eu/pbt.

[26] ECHA, "List of Substances Subject to POPs Regulation," 2024. https://echa.europa.eu/list-of-substances-subject-to-pops-regulation.

[27] UNEP, "The New POPs under the Stockholm Convention," 2024. http://www.pops.int/TheConvention/ThePOPs/TheNewPOPs/tabid/2511/Default.aspx.

[28] European Chemicals Agency, "ECHA Registered Substances," 2024. http://echa.europa.eu/information-on-chemicals/registered-substances.

[29] SMILECAS Corporation, "SMILECAS Database," 2024. http://www.srcinc.com/what-we-do/product.aspx?id=135.

[30] European Commission, EC Inventory–Institute for Health and Consumer Protection (JRC-IHCP), "EINECS," 2024. http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/information-sources/ec_inventory.

[31] Umweltbundesamt, "Klassifizierung Der PBT-Eigenschaften Der Ehemaligen Neustoffe Durch QSAR-Berechnungen Für Die Identifizierung Weiterer Besonders Besorgniserregender Stoffe Unter REACH," ETH Zurich and the German Federal Environment Agency, 2024.

[32] ECHA, "Guidance on Information Requirements and Chemical Safety Assessment Chapter R.16: Environmental," 2008.

[33] ECHA, "Guidance on Information Requirements and Chemical Safety Assessment: Chapter R.16 – Environmental," 2008.

[34] Bento, A. P.; Hersey, A.; Félix, E.; Landrum, G.; Gaulton, A.;Atkinson, F.; Bellis, L. J.; De Veij, M.; Leach, A. R. An Open Source Chemical Structure Curation Pipeline Using RDKit. J. Cheminf. 2020,12 (1), 1-16