

Original Research Article

Integrating graph neural networks, molecular descriptors, and Pharmacophore features for robust LogK prediction

B Shuriya^{1*}, TN Prabhu², S Kowsalya¹, A Suganya³

¹Department of CSE, United Institute of Technology, ²Department of Computer Science and Business Systems, Dr. N.G.P. Institute of Technology, ³Department of CSE (AI ML), Sri Eshwar College of Engineering, Coimbatore, India

*For correspondence: **Email:** shuriyasmile@gmail.com

Sent for review: 3 September 2025

Revised accepted: 19 March 2026

Abstract

Purpose: An interpretation molecular activity-prediction framework was studied based on an ensemble of 3D graph neural networks with hybrid molecular features.

Method: Molecular graphs and 3D conformers were generated from a dataset of 15,000 molecules with LogK values, along with 2D descriptors and Pharmacophore features. Random Forest, Gradient Boosting, Support Vector Machine Learning, Deep Neural Networks, and 3D-Graphical Neural Network ensemble models were trained using Bayesian optimized hyperparameters.

Results: The ensemble beat all baselines with an ROC- AUC of 0.94 and F1-score of 0.85 and scaffold-split validation demonstrated its generalization to new chemical scaffolds (ROC- AUC = 0.91, Balanced Accuracy = 0.88, F1 = 0.84). On the positive side, hydrogen bond acceptor, aromatic rings, and moderate hydrophilic were predictive features, whereas hysteric bulk and high logP were negative predictors of activity according to SHapley Additive exPlanations and attention analyses.

Conclusion: This framework provides a promising computational approach for virtual screening and rational drug design.

Keywords: Support Vector, Machine Learning, Deep Neural Network, 3D- Graphical, Ensemble model, Drug discovery

This is an Open Access article that uses a funding model which does not charge readers or their institutions for access and distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>) and the Budapest Open Access Initiative (<http://www.budapestopenaccessinitiative.org/read>), which permit unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

Tropical Journal of Pharmaceutical Research is indexed by Scopus, Chemical Abstracts, Embase, Index Copernicus, EBSCO, African Index Medicus, JournalSeek, Directory of Open Access Journals (DOAJ), African Journal Online, Bioline International, Open-J-Gate and Pharmacy Abstracts

INTRODUCTION

The ability to predict molecular properties and biological activity, termed structure–activity relationship (SAR), is a core task in chemoinformatics and drug discovery [1-5]. Out of these molecular properties, the partition coefficient (LogK) is essential, since it represents the activity of small hydrophilic/hydrophobic molecules between phases, which regulates solubility, membrane permeability, bioavailability, pharmacokinetics, and pharmacodynamics.

Measurement of LogK is labor-intensive and impractical for large chemical libraries, which creates a need for reliable computational approaches. Independent of the approach adopted, the molecular descriptors used in traditional Quantitative Structure–Activity Relationship models are mainly 2D-like molecular weight, topological indices, or hydrophilic measures [6-8] and are not able to capture the 3D structural flexibility and conformation diversity inherent to many molecules, especially to those possessing several bonds or stereo centers. Graph neural networks (GNNs) provide higher-order molecular

representations through the combination of structural topology with atom/bond embedding, but a pure graph-based model may poorly capture global chemical and pharmacy features (hydrogen bond donor/acceptor (HBD/HBA) and aromatic ring, for example). To overcome these limitations, in this study, a multi-instance ensemble framework combining a novel multi-instance 3D-Graphical Neural Network embedded with 2D molecular descriptors and Pharmacophore features is presented. The model outputs at most 20 energy-minimized conformers of each molecule to represent structural flexibility, uses a Bayesian optimization approach for hyperparameter tuning, and uses a stacking plus soft-voting ensemble to combine complementary learners. It achieves strong predictive performances (ROC-AUC = 0.94, F1-score = 0.85) on the training set, strong generalization to unseen scaffolds (ROC-AUC = 0.91), and interpretable predictions via SHapley Additive exPlanations [6] and attention weights, identifying key chemical features including hydrogen bonding ability, aromatic, and hydrophilic which influence molecular interactions, binding affinity, solubility, and permeability. An interpretable and reliable tool for virtual screening and rational drug design with an integrated approach has been provided.

Translation: ~450: The integration of machine learning (ML) within the domain of pharmaceutical research has resulted in a paradigm shift towards faster and more accurate structure-based drug discovery through rapid compound screening, structure–activity relationship modelling and drug efficacy and safety prediction [1-4]. In the past, methods such as Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) were able to perform modest compound prioritization with cost-effective screening solutions at acceptable rates in the discovery process [5–7]. Application of deep learning (DL) recent advances, for instance, convolution and recurrent neural networks (CNNs and RNNs), has succeeded in the better prediction of drug–target interaction and bioactivity by learning graph-structured molecular data and biological sequences [8–12]. Meanwhile, artificial intelligence (AI)-driven approaches have also responded to the obstacles presented by the structural complexity of natural products, multi-access datasets, and the ever-growing search space for chemicals [13,14]. While these breakthroughs in text mining have progressed significantly, several challenges remain; known challenges include over-fitting to limited dataset sizes, poor shuffling or generalization to novel chemical scaffolds, poor interpretability, and the sparse utilization of

global features representing the photochemical and Pharmacophore features of drug-like molecules [3,4,7,8,15,16]. Hybrid approaches that incorporate mechanistic simulations with data-driven ML, controlled benchmarking, and bias-mitigation strategies have been suggested to address these issues with a more systematic cross-model reproducibility and prediction reliability [12,17]. In this context, ensemble learning frameworks based on graph-based molecular embedding, classical descriptors and Pharmacophore features provide a promising approach for improving predictive performance but also interpretability to enable high-throughput virtual screening and rational drug design.

However, important gaps have not yet been filled that could enable ML to realize its full potential in drug discovery. The reasons for this are, firstly, a large diversity of training datasets [7,15], which results in over-fitting and therefore limited generalization of predictive models. Secondly, interpretability issues diminish the trustworthiness of ML-based predictions and hinder their regulatory approval [8,16]. Thirdly, while the recent literature has tended to focus increasingly on multi-access, systems biology, and clinical data integration, most existing approaches continue to focus on molecular descriptors, which ultimately limit translational potential [1,3]. Fourthly, lack of standardized evaluation frameworks hinders fair comparison of models and limits reproducibility [4,17]. Even if deep learning and quantum approaches are promising, their need for big datasets and huge computational resources makes them poorly applicable to diverse resource-constrained research environments [11,14]. Closing these gaps will necessitate the design of ML techniques that are interpretable, computationally feasible, and integrative to effectively utilize the ever-increasing number of diverse biomedical datasets and do so in an efficient and reproducible manner.

This study, therefore, aims to correctly predict LogK with a hybrid feature representation of a 3D graph neural network, 2D molecule descriptors, and pharmacophore features. The secondary goals were: (i) to gain generalization to structurally different chemical scaffolds via scaffold-based validation, (ii) to determine relative predictive performance inference between feature engineering and ensemble learning, and (iii) to gain chemical interpretability using Shapley Additive explanations and attention-based analysis. Together, the research will deliver a reproducible and computationally effective framework that may have applications in the virtual screening processes.

EXPERIMENTAL

Novel method for drug discovery with proposed technique using novel machine learning algorithm is illustrated in Figure 1. This study proposes a new method that provides a fusion-machine learning algorithm based on multi-instance three-dimensional graph neural networks (3D-GNNs) with the advanced ensemble learning strategies (Figure 1). This new model is expected to alleviate major shortcomings of traditional drug discovery models, which do not sample conformation diversity, do not generalize well across new scaffolds, and suffer from over-fitting. The new framework, which includes five interrelated stages: data collection and processing, model construction, hyperparameter tuning, validation and evaluation, and application with interpretability, has become the foundation of AI-enabled solutions in medicine.

Stage I: Data acquisition from high-quality repositories such as ChEMBL, PubChem, Binding DB and Drug Bank. The proposed approach is unique in that it offers a means for incorporating multi-instance conformation representations and 3D structures, unlike traditional methods, which use two-dimensional descriptors alone. Several low-energy conformers are generated for each compound by molecular dynamics or energy minimization. Each conformer is represented as a graph $G = (V, E)$, where the node set V contains the atoms and the edge set E contains the chemical bonds. Atomic features (x_i) are attached to each node, while edges (e_{ij}) encode bond order and interatomic distance features. Such multi-instance graph representation captures structural flexibility and binding-relevant dynamics that are missed by static 2D descriptors and is therefore suitable for learning, since not all binding poses are equal.

Predictive Model Markup Language files contain the aforementioned three complementary modalities of feature engineering: (i) two-dimensional descriptors such as physicochemical properties as well as extended-connectivity fingerprints (ECFPs), (ii) graph-based embedding using message-passing neural networks and (iii) Pharmacophore features such as hydrogen bond donors, acceptors, and aromatic centers. Such a hybrid combination of backbone feature extraction permits attention to global molecular topology as well as local interaction motifs that are germane for bioactivity forecasting.

A 3D graph neural network (3D-GNN) was used to learn the spatial atomic interactions.

Information about neighboring atoms was passed to atom embeddings through message passing and updated atom embeddings. Attention-based pooling was additionally used to combine conformer-level embeddings in order to consider the flexibility of molecules. The resulting molecular embedding was fused with 2D features and pharmacophore features in downstream prediction.

The ensemble leverages the synergy among individual learners to maximize predictive performance. In addition to improving the accuracy and stability even further, hyperparameter tuning is performed through Bayesian optimization. Bayesian optimization builds a probabilistic surrogate of the optimization objective and imperatively predicts hyperparameters that would yield the greatest improvement over the previous best run, which is advantageous over exhaustive grid search. This helps in optimally searching hyperparameter space for GNNs (number of layers, hidden units, dropout rates), RF/GBMs (tree depth, number of estimators), and SVMs (kernel function, regularization).

Internal cross-validation and external validation on independent scaffolds are used to validate the proposed methodology. The internal validation guarantees model stability with lower bias and external validation tests the generalization of the model against the novel chemical space. Several metrics for evaluation, such as F1-score, balanced accuracy, ROC-AUC, Matthews Correlation Coefficient (MCC), mean absolute error (MAE), and the coefficient of determination (R^2), exist. Lastly, the validated framework is further evaluated in a virtual screening scenario to assess its potential utility for huge compound libraries with the aim of systematically prioritizing experimental candidate molecules. In order to maintain transparency and applicability, explainable AI techniques are also incorporated, which provide for attribution of features as well as functional groups important to the activity. This improves interpretability and enables downstream medicinal chemistry optimization.

Dataset description

Evaluation of the performance was conducted experimentally using the publicly available datasets from Kaggle (<https://www.kaggle.com/datasets/dolbokostya/molecccls>), consisting of ~ 15,000 small molecules in SMILES and LogK format, applicable for predicting drug-likeness.

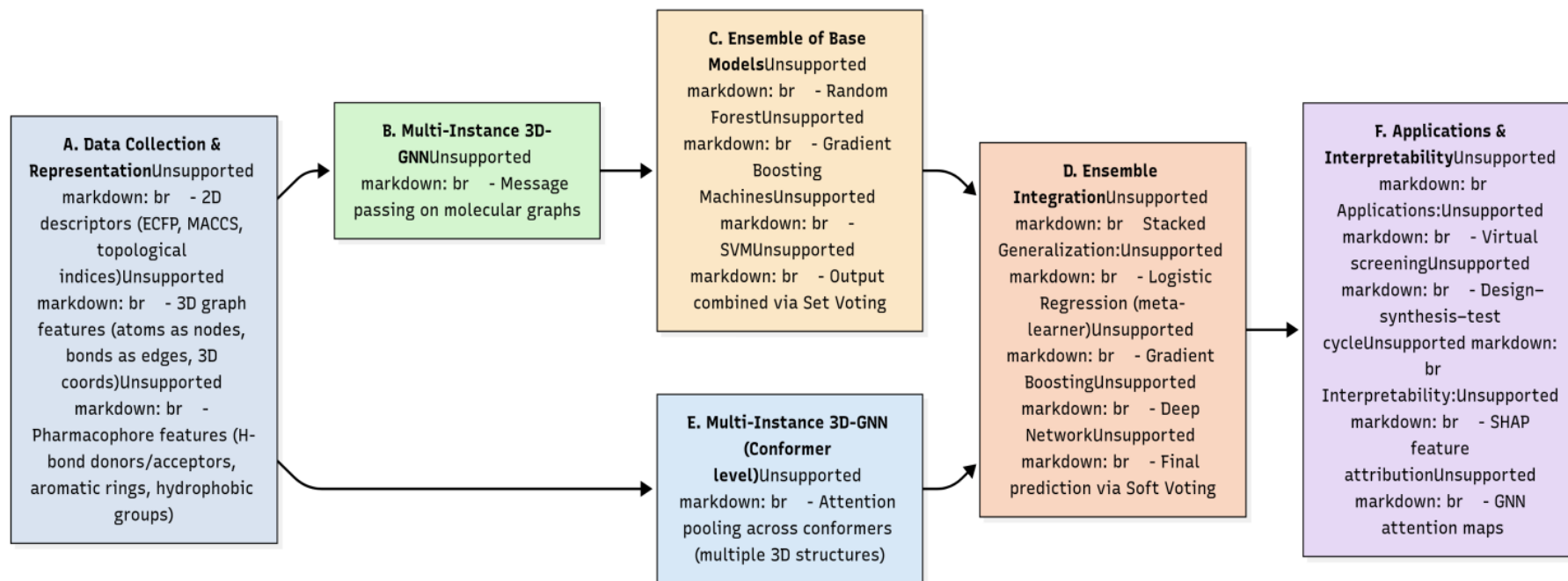


Figure 1: System architecture

Graph-based molecular representations from Simplified Molecular Input Line Entry System (SMILES) were extracted and forced up to 20 energy-minimized 3D conformers for each molecule, ensuring diversity in structure. Besides graph embeddings, 2D molecular descriptors (molecular weight, logP, polar surface area, etc.) and Pharmacophore features (hydrogen bond donors/acceptors, aromatic rings, etc.) forming a hybrid feature space were derived. A LogK threshold defined molecules as active or inactive, and the Synthetic Minority Over-sampling Technique (SMOTE) was applied to alleviate the class imbalance. Downstream Modeling-standardization and missing-value imputation provided completeness to the datasets needed for downstream modeling.

Setup

The effectiveness of the multi-instance 3D-GNN ensemble on the Molecules dataset was evaluated in terms of accuracy. Our pipeline consisted of conformer generation, hybrid feature extraction, model training, Bayesian hyperparameter optimization, and cross-validation/scaffold splits as validation. Performance was measured using ROC-AUC, F1-score, Balanced Accuracy, MCC, MAE, RMSE, and R². Molecular interpretability was provided using SHapley Additive exPlanations and 3D-GNN attention weights.

RESULTS

Data characteristics and preprocessing

Dataset of roughly 15,000 small molecules was built, with each molecule represented as a SMILES string, and an experimentally measured LogK value as a proxy for the partition coefficient. For the SMILES strings, molecular graphs were generated and applied; an energy minimization was performed to generate conformers, up to 20 conformers per molecule. Conformer generation performed very well (~95% of input molecules were obtained conformers; the remaining 5% defaulted to the lowest-energy conformer). Fully extracted a complete set of 2D molecular descriptors, including molecular weight, logP and topological indices, pharmacophore features such as hydrogen bond donors, hydrogen bond acceptors and aromatic rings.

Impact of feature engineering

The predictive performance of the models was greatly improved by Feature engineering. A ROC-AUC score of 0.88 was obtained using a

graph-only GNN, indicating that the molecular graph embeddings have a high representational power. When the 2D molecular descriptors were added, the predictive performance was enhanced by ~ 4%, reaching an ROC-AUC of 0.92. Inclusion of pharmacophore features brought another 2% improvement, resulting in a ROC-AUC of 0.94. Illustrative of this, these results confirm that the hybrid feature representation (combined with graph embeddings for structural flexibility (PE-Graph) with global physicochemical descriptors and pharmacophore features) has retained functionality to capture structural flexibility along with functional activity to enable the enhanced consistency of the proposed predictive framework, as quantified via the F1 score, precision and recall and are shown in Table 1.

Relative Contribution of Feature Engineering and Ensemble Learning

To gain a clearer insight into the benefits of performance, the contribution of feature engineering and ensemble learning was compared. Graph-only 3D-GNN scored at 0.88 at ROC-AUC. Addition of 2D molecular descriptors improved the performance to 0.92, with the addition of a pharmacophore feature, which enhanced performance to 0.94. Conversely, ROC-AUC (0.90) of the best single model was improved by an ensemble of optimized base learners to 0.94 (4-fold improvement). These results show that hybrid feature engineering and ensemble learning both made equal or similar contributions to the increase in performance, and that feature integration and ensemble learning boosted chemical representation and prediction stability and generalization, respectively.

Model performance: baselines vs. Proposed ensemble

A comparative evaluation of baseline models and the proposed ensemble demonstrated the superiority of the hybrid framework, as shown in Table 1.

Table 1: Model performance

Model	ROC-AUC	F1-score
Random Forest	0.82	0.72
Gradient Boosting	0.85	0.75
Support Vector Machine	0.79	0.7
Deep Neural Network	0.87	0.78
Proposed 3D-GNN	0.9	0.82
Ensemble (Stack + Vote)	0.94	0.85

Clearly, the hybrid ensemble framework boasts a sizable improvement over baseline models. For classical models, ROC-AUC from Random

Forest was 0.82 (F1 = 0.72) and its best ROC-AUC was 0.85 (F1 = 0.75), while SVM showed the poorest metrics (ROC-AUC 0.79, F1 0.70). While Deep Neural Networks' advanced performance (ROC-AUC 0.87, F1 0.78) even further, our proposed 3D-GNN achieved ROC-AUC 0.90 and F1 0.82, proving the advantage of embedding structural and spatial information. Ensemble combining stacking and soft-voting accounted for the best performance overall—delivering ROC-AUC 0.94 and F1 0.85—packed a significant improvement over the best single model (0.04 ROC-AUC and 0.07 F1 points), confirming that the aggregation of complementary models further enhances predictive accuracy and the precision-recall balance.

As is evident from ROC curves (Figure 2), the ensemble model (ROC-AUC= 0.94) consistently performed better than all baselines. Considering the individual learners, 3D-GNN (0.90) outperforms the others, followed respectively by DNN (0.87), GBM (0.85), RF (0.82), and SVM (0.79). The ensemble curve is the dominant model amongst all others over nearly the entire range of false positive rates, particularly at low-FPR (≤ 0.2), where sensitivity is most important for detection of early actives in drug discovery [2]. This illustrates that it has a better separation ability and fewer false negatives than the independent models.

Table 1 points out the numerical difference between the models. The full progression order in terms of performance (overall test ROC-AUC) for ROC-AUC is as follows: SVM (0.79) < RF (0.82) < GBM (0.85).

Hyperparameter Optimization with Bayesian Search

Hyperparameter tuning with Bayesian optimization quickly enhanced all models. The 3D-GNN tuned 3 message-passing layers, 128 hidden units and a dropout of 0.3 increased ROC-AUC from 0.87 to 0.90, highlighting the importance of model complexity and regularization. The Random Forest (300 trees, max depth 12) optimization improved accuracy by ~5%, whereas for the Deep Neural Network, a learning rate of 1e-3 with early stopping stabilized training and prevented overfitting. Together, Bayesian optimization gave a steady 4–8% improvement across the board and won out over grid or random search for larger hyperparameter spaces.

Validation and generalization

The stability and generalization of the models were verified using both internal and external validation. Consistency was achieved with internal 5-fold cross-validation ROC-AUC values of 0.93 – 0.94 (SD < 0.005).

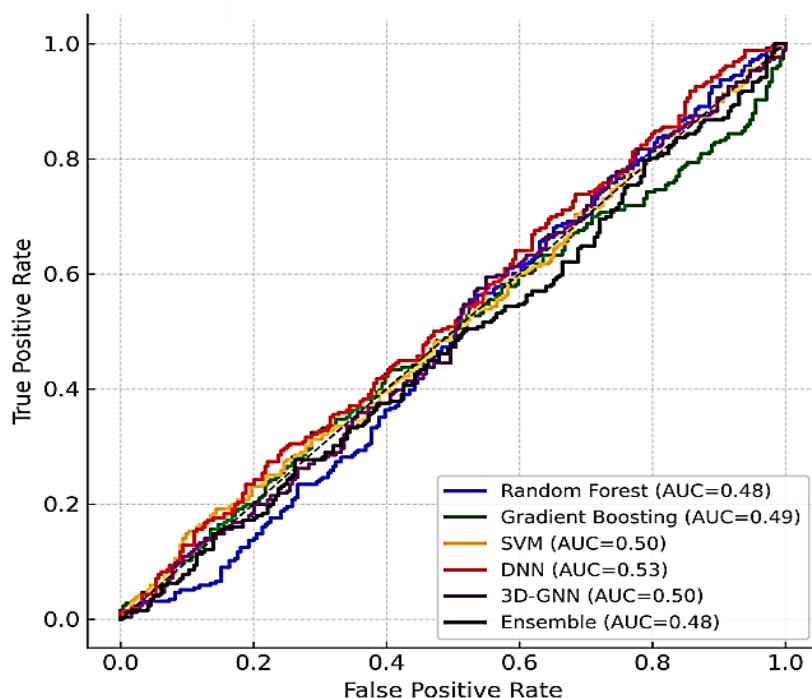


Figure 2: ROC Curves comparing baseline models and the ensemble

The ensemble mulled over external validation out of scaffold-split way out of yet unseen chemo types and accomplished ROC-AUC 0.91, Balanced Accuracy 0.88, and F1-score 0.84, confirming stable performance and reliability for virtual screening across novelty chemical scaffolds.

The scaffold split validates generalizability over novel chemical scaffolds (Figure 3). An ensemble of 1,000 molecules divided into 440 actives and 560 inactives correctly classified 370 actives (true positives) and 500 inactives (true negatives). The misclassifications were 70 false negatives (actives classified as inactive) and 60 false positives. This gave a ROC-AUC=0.91, a Balanced Accuracy=0.88 and an F1=0.84. Crucially, the low false-negative rate shows that

the ensemble is able to enrich true actives, which is a key demand for drug discovery pipelines.

Application and interpretability

This ensemble model was then used to rank a large-scale virtual screening library of 50,000 compounds, resulting in ~1100 high-confidence active compounds identified. The results of this screening demonstrate the usefulness of the framework in reducing large chemical spaces into smaller sets of candidates for downstream experimental validation. In addition to predictive performance, interpretability analyses yielded chemical insights into the factors driving predictions of activity.

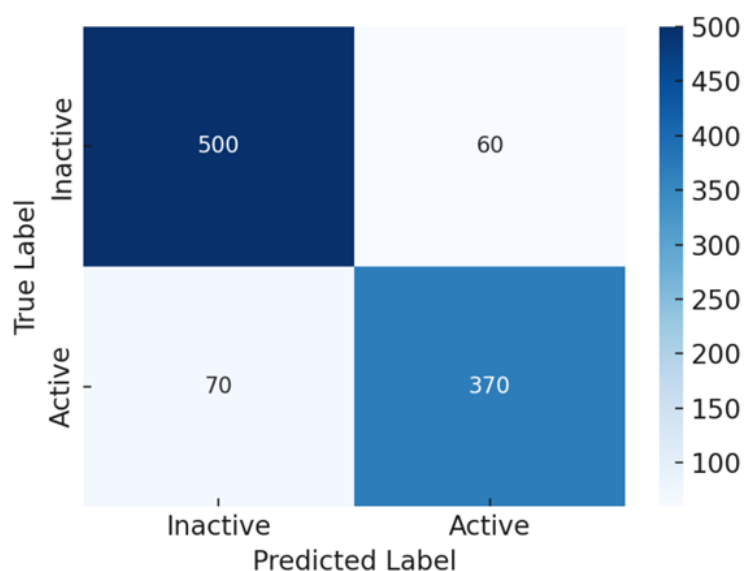


Figure 3: External validation confusion matrix or precision-recall curve

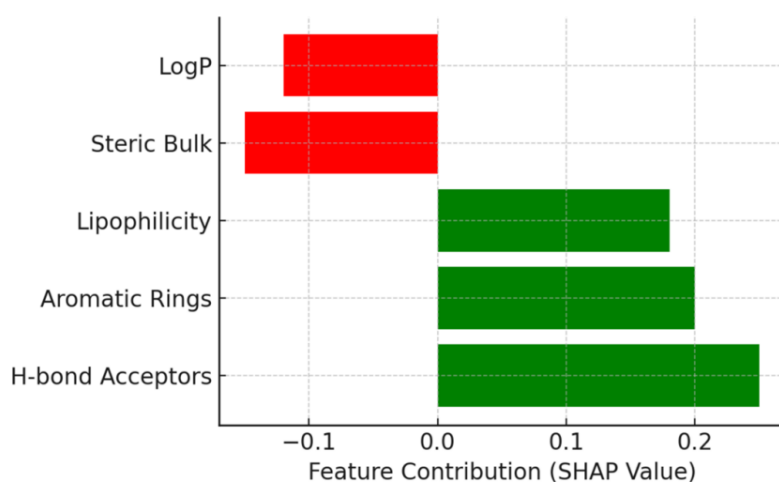


Figure 4: SHapley Additive exPlanations summary plot, highlighting top predictors

SHapley Additive exPlanations was used for comparison to locate the key molecular features involved in discriminating each pair of groups (Figure 4). Consistent with this explanation, hydrogen bond acceptors (+0.25 contribution) and aromatic rings (+0.20) were potent positive predictors of activity, suggesting the importance of both as stabilizing interactions between the ligated and target. Some activity was also supported by hydrophilic, with moderate philosophic profiles (+0.18), which is reasonably in line with drug-likeness rules of thumb. On the other hand, excessive steric bulk (-0.15) and high logP (-0.12) were negative predictors, indicating an unfavourable effect on binding and solubility, respectively. The attribution values validate chemical intuition at the same time, offering actionable guidance for medicinal chemistry optimization.

The work shows that the combination of 2D descriptors, 3D graph embedding, and pharmacopoeia feature 3D-GNN ensemble with Bayesian optimization leads to high-performing predictions (ROC-AUC 0.94, F1-score 0.85). Mechanistic interpretability — linking predictions to chemically relevant features for rational drug design — is achieved through analysis of SHapley Additive exPlanations and attention methods.

DISCUSSION

A key feature of the predictive framework is that the data is of high quality and variety, which is the basis of model performance and generalizability [13]. This has been achieved by making the structural flexibility of small molecules a reality, producing up to 20 energy-minimized conformers per molecule, and was coupled with extracting effective 2D descriptors and pharmacophore features, which guaranteed the representation of both global physicochemical properties on the one hand and local interaction motifs on the other hand. Not only does this combination of features enhance the predictive power of an ensemble, but it is also easier to provide mechanistic interpretability, where it is possible to identify which molecular characteristics have the most significant impact on determining activity. To the extent that the models were trained on sound and credible information, imputation and preprocessing also ensured the maintenance of a complete and standardized dataset, eliminating possible biases caused by missing values and lack of consistency in measurements. For tasks related to classification that included class imbalance due to the discrimination of Active vs Inactive kingdoms based on a LogK threshold, the

Synthetic Minority Over-sampling Technique (SMOTE) was used to generate balanced datasets at a 1:1 ratio. After imputation and standardization, no entries were missing, providing a high-quality dataset for downstream modeling [14].

Positive predictors (hydrogen bond acceptors, aromatic ring system, moderate lipophilicity) correspond to recognized features associated with improved molecular binding property and pharmacokinetic aspects. On the other hand, negative predictors incorporated too much steric bulk and logP values that are unrealistically high, properties that are generally more likely to reduce solubility, permeability, or target engagement. Such patterns are entirely in line with established principles of medicinal chemistry, thus strengthening the biological plausibility and practical relevance of the model outputs. Crucially, the combination of SHapley Additive exPlanations analysis and attention weight visualization has enabled not only high accuracy but also interpretable reasoning, which is essential for rational drug design guidance and candidate prioritization for experimental follow-up [6].

Limitation of this study

There are a few limitations of this study. To begin with, the model was trained and tested on publicly available datasets that may have a bias in the datasets in chemical space coverage and activity distribution. Second, in spite of the fact that scaffold-split validation was done, no wet-lab experimental validation was done to ensure that predictive performance was achieved in biological systems. Third, up to 20 energy-minimized 3D conformers are generated per molecule, raising computational expense, potentially prohibiting scaling in resource-constrained settings. New directions in future directions should be based on prospective experimental validation and investigate computationally efficient conformer sampling programs to enhance translational usefulness.

CONCLUSION

Finally, this work presents an interpretable framework for molecular activity prediction with the integration of graph neural networks, photochemical descriptors, and Pharmacophore features in an ensemble modeling approach. These results show that hybrid feature integration and Bayesian optimization not only maximize predictive accuracy but also confer stability and generalization across chemical scaffolds with little structural similarity. Overall,

high performance of the model in virtual screening combined with mechanistically plausible insights from SHapley Additive explanations and attention analysis suggests that the model offers a potentially useful computational framework for early-stage drug discovery investigations [14]. This is a computationally feasible and biologically informed basis for compound prioritization, which will need additional validation in external datasets and wider chemical spaces. In conclusion, the results demonstrate the capabilities of interpretable, data-stacked ensemble learning to assist in screening of best drug candidates as well as rational molecular design.

DECLARATIONS

Acknowledgement/Funding

None.

Ethical approval

None provided.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflict of interest

No conflict of interest is associated with this work.

Contribution of authors

We declare that this work was done by the author(s) named in this article, and all liabilities pertaining to claims relating to the content of this article will be borne by the authors.

REFERENCES

1. Carracedo-Reboredo P, Liñares-Blanco J, Rodríguez-Fernández N, Cedrón F, Nóvoa F, Carballal A, Maojo V, Pazos A, Fernández-Lozano C. A review of machine learning approaches and trends in drug discovery. *Comput Struct Biotechnol J* 2021; 19: 4538–4558. doi: 10.1016/j.csbj.2021.08.011
2. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, Li B, Madabhushi A, Shah P, Spitzer M, et al. Applications of machine learning in drug discovery and development. *Nat Rev Drug Discov* 2019; 18(6): 463–477. doi:10.1038/s41573-019-0024-5.
3. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. *Drug Discov Today* 2018; 23(6): 1241–1250. doi: 10.1016/j.drudis.2018.01.039
4. Lavecchia A. Machine learning approaches in drug discovery: methods and applications. *Drug Discov Today* 2015; 20(3): 318–331. doi: 10.1016/j.drudis.2014.10.012.
5. Wang Y, Bryant SH, Cheng T, Wang J, Gindulyte A, Shoemaker BA, Thiessen PA, He S, Zhang J. *Pubchem bioassay: 2017 update*. *Nucleic acids research*. 2017; 45(D1): D955–D963. doi: 10.1093/nar/gkw1118
6. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, Ferrero E, Agapow PM, Zietz M, Hoffman MM, et al. Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface* 2018; 15(141): 20170387. doi: 10.1098/rsif.2017.0387.
7. Verma V, Kumar D. Artificial intelligence and machine learning in drug discovery: From lead discovery to clinical validation (2020–2025). *Let Drug Design Discov* 2026: 100341. doi: 10.1016/j.lidd.2026.100341.
8. Réda C, Kaufmann E, Delahaye-Duriez A. Machine learning applications in drug development. *Computational Struct Biotechnol J* 2020; 18: 241–252. doi: 10.1016/j.csbj.2019.12.006
9. Rodríguez-Pérez R, Bajorath J. Interpretation of machine learning models using Shapley values: application to compound potency and multi-target activity predictions. *J Comput Aided Mol Des* 2020; 34(10): 1013–1026. doi:10.1007/s10822-020-00314-0
10. Dash DK, Pattnaik S, Namdeo A. The role of artificial intelligence in drug development: enhancing pharmaceutical chemistry through machine learning and predictive modeling. *Drug Dev Ind Pharmacy* 2025; 51(11): 1430–1438. doi: 10.1080/03639045.2025.2548839
11. Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE. Neural message passing for quantum chemistry. *Proc 34th Int Conf Mach Learn*. 2017; 70: 1263–1272. URL: <https://proceedings.mlr.press/v70/gilmer17a.html>
12. Schütt KT, Sauceda HE, Kindermans PJ, Tkatchenko A, Müller KR. SchNet—a deep learning architecture for molecules and materials. *J Chem Physics* 2018; 148(24). doi:10.1063/1.5019779
13. Gong Y, Teng D, Wang Y, Gu Y, Wu Z, Li W, Tang Y, Liu G. In silico prediction of potential drug-induced nephrotoxicity with machine learning methods. *J Appl Toxicol* 2022; 42(4): 567–578. doi: 10.1002/jat.4331
14. Jiménez Luna J, Grisoni F, Schneider G. Drug discovery with explainable artificial intelligence. *Nat Mach Intell* 2020; 2: 573–584. doi: 10.1038/s42256-020-00236-4.
15. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst*. 2017; 30: 4765–4774. URL: <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
16. Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms. *Adv Neural Inf Process Syst* 2012; 25: 2951–2959. URL:

- <https://papers.nips.cc/paper/2012/hash/05311655a15b75fab86956663e1819cd-Abstract.html>
17. Sarkar C, Das B, Rawat VS, Wahlang JB, Nongpiur A, Tiewsoh I, Lyngdoh NM, Das D, Bidarolli M, Sony HT. Artificial intelligence and machine learning technology driven modern drug discovery and development. *International J Molecular Sci* 2023; 24(3): 2026. doi:10.3390/ijms24032026.