hERG-LTN: A New Paradigm for hERG Cardiotoxicity Prediction Using Neuro-Symbolic and Generative AI

Delower Hossain, Fuad Al Abir, Jake Y Chen*

Abstract— **Assessing adverse drug reactions (ADRs) during drug development is essential for ensuring the safety of new compounds. The blockade of the Ether-à-go-go-related gene (hERG) channel plays a critical role in cardiac repolarization. Computational predictions of hERG inhibition can help foresee drug safety, but current data-driven approaches have limitations. Therefore, a new paradigm that bridges the gap between data and knowledge offers an alternative for advancing precision pharmacogenomics in assessing hERG cardiotoxicity. This study aims to develop a reasoning-based, in-silico, robust model for predicting drug-induced hERG inhibition, facilitating new drug development by reducing time and cost, and supporting downstream in-vitro and in-vivo testing. In this study, we constructed a new dataset (UnihERG_DB) by sourcing data from ChEMBL, PubChem, BindingDB, hERG Karim's, hERG Blocker's, and GTP bioactivity databases. The final dataset comprises 20,409 SMILES (Simplified Molecular Input Line Entry System) data samples, labeled as hERG blockers (IC50 < 10 μM) or non-hERG blockers (IC50 ≥ 10 μM). Molecular features were extracted using Morgan and CDK fingerprints. Furthermore, we explored embedding feature computation using cutting-edge Large Language Models, including NVIDIA MegaMolBART, LLaMA 3.2, Gemini, and DeepSeek. Finally, we utilized the Logic Tensor Network (LTN), an advanced AI framework, to train and develop the hERG predictive model. Model performance was evaluated using two benchmarks: External Test-1 and hERG-70. The model outperformed several SOTA approaches, including CardioTox, M-PNN, DeepHIT, CardPred, OCHEM Predictor-II, Pred-hERG 4.2, Random Forest, and Gradient Boosting. On the External Test-1 dataset, hERG-LTN achieved an accuracy (ACC) of 0.931, a specificity (SPE) of 0.928, and a sensitivity (SEN) of 0.933. Furthermore, on the hERG-70 benchmark, our method achieved an accuracy of 0.827, a specificity of 0.890, and a correct classification rate (CCR) of 0.833. Overall, the Neuro-Symbolic AI approach sets a new standard for hERG-related cardiotoxicity assessment, yielding competitive results with current state-of-the-art (SOTA) models, and highlights its potential for advancing precision pharmacogenomics in drug discovery and development (GitHub: https://github.com/hossain013/hERG-LTN).**

Keywords— hERG; Neuro-Symbolic AI; Drug Discovery; Cardiotoxicity; Machine Learning; LLM Embeddings.

I. INTRODUCTION

The study of Cardiotoxicity and its impact on drug discovery has become a spirited area of research in recent years. The human Ether-à-go-go-Related Gene (hERG) encodes a potassium channel essential for cardiac repolarization. Drugs that block this channel can lead to prolonged Q.T. intervals on an electrocardiogram (ECG), increasing the risk of arrhythmias as well as sudden cardiac death. Therefore, accurately classifying this type of Adverse Drug Reactions (ADRs) is pivotal during drug development. Several studies have leveraged A.I. methodologies to assess Drug-Induced Cardiotoxicity (DICT) [1]-[3]. Our studies have identified various QSAR (Quantitative Structure-Activity Relationship) models [4]-[5] that employed machine and deep learning strategies [6]-[20], including auto-QSAR modeling, to predict hERG inhibition. Current state-of-the-art models include CToxPred-hERG [21], CardioTox [22], M-PNN [22], DeepHIT [24], CardPred [25], ADMETlab 2.0 [26], ADMETsar 2.0 [27], OCHEM Predictor-I and II [28], Pred-hERG 4.2 [29], and CardioGenAI [30].

Unlike traditional predictive approaches that rely solely on standard molecular properties or conventional AI techniques, integrating domain knowledge with data is decisive in healthcare, particularly where reasoning, explainability, and interpretability are vital. However, this integration remains underexplored in hERG inhibition prediction research. In recent years, Neuro-symbolic (NeSy) approaches [31], a new dimension of AI, have garnered significant attention for combining neural networks' strengths with symbolic reasoning, resulting in more interpretable and insightful predictions. Our study discovered several innovative NeSy models across both healthcare [58] and non-healthcare domains, including (Gene research) KBANN [32], (Diabetic Retinopathy) ExplainDR [33], (Link Prediction) NeuralLP [34], (Ontology) RRN [35], NSRL [36], Neuro-Fuzzy [37], FSKBANN [38], DeepMiRGO [39], NS-VQA [40], DFOL-VQA [41], LNN [42], NofM [43], PP-DKL [44], FSD [45], CORGI [46], NeurASP [47], XNMs [48], Semantic Loss [49], NS-CL [50], LTN [51], and NESYDPP4-QSAR [59]. Moreover, the advent of generative AI is poised to revolutionize the precision pharmacogenomics industry by accelerating drug discovery, particularly in chemical compound research, and utilizing their embedding features presents a competitive performance in Cheminformatics.

This study introduces hERG-LTN, a new paradigm that integrates data-driven and knowledge-based techniques (Neuro-Symbolic AI) to enhance in-silico cardiac toxicity assessment. The model combines generative AI-based molecular embeddings with traditional structural features such

* Corresponding Author.

Jake Y. Chen is the Founding Director, Systems Pharmacology AI Research Center (SPARC), Triton Endowed Professor of Biomedical Informatics and Data Science, Adjunct Professor of Genetics, Computer Science, and Biomedical Engineering at The University of Alabama at Birmingham, AL, 35294, USA (e-mail: jakechen@uab.edu).

Delower Hossain is a 3rd year PhD student in the Department of Computer Science at The University of Alabama at Birmingham, AL, 35294, USA. (e-mail: mhossai5@uab.edu).

Fuad Al Abir is a 2nd year PhD student in the Department of Computer Science at The University of Alabama at Birmingham, AL, 35294, USA. (e-mail: fuad021@uab.edu).

as Morgan and CDK fingerprints. To ensure robustness and diversity, we constructed a large-scale dataset by aggregating bioactivity sample records from ChEMBL [53], PubChem BindingDB [54], hERG Karim's dataset [22], hERG Blockers [57], and the GTP [56]. The hERG-LTN was built with Logic Tensor Network (LTN) [51], a neuro-symbolic approach that combines the data-driven learning capabilities of neural networks with the reasoning power of first-order logic. A comprehensive experiment was conducted comparing molecular representation strategies, including fingerprints, descriptors, and embedding features generated by NVIDIA's MegaMolBART [52], LLaMA 3.2 [61], Gemini [62], and DeepSeek-QWEN_1.5b [63]. The hERG-LTN model outperformed current state-of-the-art methods on the External Test-1 dataset and demonstrated superior results on the hERG-70 benchmark, setting a new performance standard in cardiotoxicity assessment.

II. Materials and Methods

This section outlines the procedures employed in experiments to assess the performance of our proposed hERG-LTN approach. The hERG-LTN paradigm leverages several features engineering strategies, including cutting-edge LLM embeddings generated by the MegaMolBART, LLaMA 3.2, Gemini, and DeepSeek-QWEN_1.5b model. This section covers the entire pipeline, including materials, data preprocessing, feature extraction, network architecture, LTN knowledge-based settings, the training and inference phases, and the system of measurements used to evaluate the model's performance against state-of-the-art models (Fig. 1).

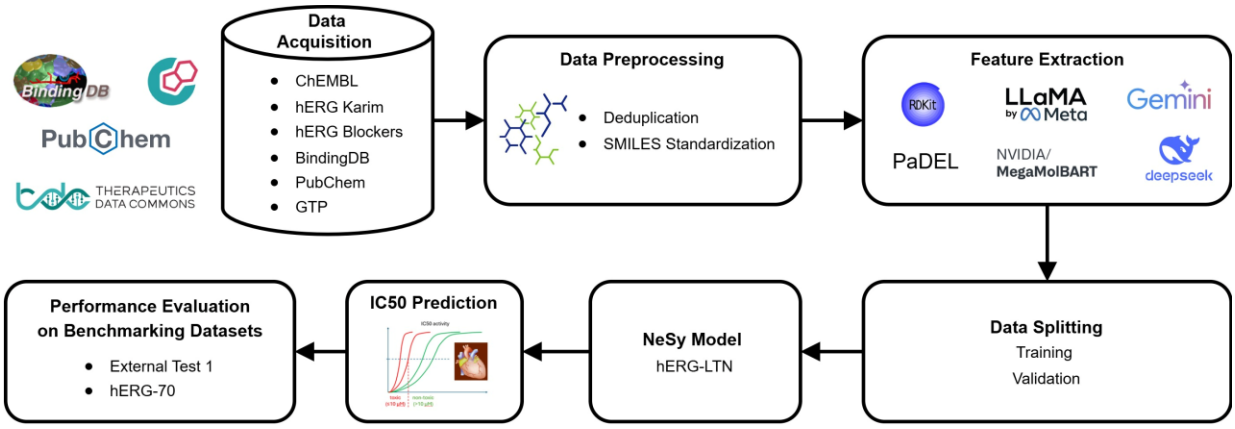

Fig. 1: Workflow of hERG-LTN Model

A. *Data acquisition*

In this study, a new cohort was constructed, UnihERG_DB, by sourcing data from several bioactivity databases: ChEMBL [53], hERG Karim's (TDC) [22], BindingDB [54], PubChem [55], Guide to Pharmacology (GTP)[56], and hERG Blockers [57]. The ChEMBL data was obtained using ID CHEMBL240, BindingDB via this link, PubChem in CSV format with NCBI Gene 3757 (link), and GTP via the corresponding link. The hERG Karim's and hERG Blocker's datasets were collected from the TDC website using the Python package PyTDC.

Table 1: UnihERG_DB Data Distributions

Source	Blocks	Non-blocks	Both
ChEMBL [53]	5460	6320	11780
hERG Karim [22]	2628	2469	5097
BindingDB [54]	952	955	1907
PubChem [55]	719	749	1468
hERG Blockers [57]	65	88	153
GTP [56]	2	2	4
Total	**9826**	**10583**	**20409**

After curation, the UnihERG_DB database consists of 20,409 trainable compounds (Table 1). The final dataset comprises 20,409 structures represented as SMILES (Simplified Molecular Input Line Entry System). labeled as hERG blockers (IC50 < 10 µM) or non-hERG blockers (IC50 ≥ 10 µM).

B. *Data Preprocessing*

Both hERG Karim and hERG Blocker's datasets offered preprocessed molecular data and corresponding binary IC50 information representing blockers (IC50<10 µM) and non-blockers (IC50>=10 µM). However, numerical IC50 measurements in nM were given in ChEMBL, BindingDB, and GTP, but those in µM were given in PubChem. We harmonized all units into Micromolar (µM), followed by thresholding as blockers (IC50 <10 µM) or non-blockers (IC50>=10 µM). We used MolVS Python-Package to standardize SMILES and applied a standard scaler by scikit-learn for embedding features. Finally, we dropped duplicates and NaN values.

C. *Feature extraction*

PadelPy [60] is a widely used tool for converting SMILES strings into numerical descriptors, offering more than 10 types of descriptors, including CDK, CDKextended, AtomPairs2DCount, AtomPairs2D, EState, CDKgraphonly, KlekotaRothCount, KlekotaRoth, MACCS, SubstructureCount, and Substructure. Each descriptor type provides different features, with CDK (1024 features) consistently delivering superior results. In this study, we extracted descriptors, fingerprints, and embedding. Fingerprints were computed using CDK PadelPy github repo,

and Morgan descriptors (nBits=2048 and radius=2) [64] were calculated via RDKit. Finally, embeddings were generated by NVIDIA's advanced LLM model, MegaMolBART [52], by following the instructions on GitHub to configure the necessary environment in Docker. In addition, we computed LLaMA 3.2, Gemini, and DeepSeek-QWEN_1.5b embeddings using hugging face models. We integrated LLM embeddings (e.g., MegaMolBART, LLaMA 3.2, Gemini, DeepSeek-QWEN) since they trained on massive data that may capture nuanced biochemical and functional relationships within SMILES sequences. The assumption was this enriches representations beyond traditional fingerprints and enhances the learning of complex cardiotoxicity molecular patterns.

D. Model Architecture

The LTN [51] framework was employed to build the hERG-LTN classifier. The merit of this approach is its capability to tackle the limitations of traditional deep learning systems, which often struggle with tasks requiring reasoning, interpretability, symbolic manipulation, and knowledge integration. The LTN architecture consists of two key components: a logical component and a neural network. The visual architecture of the classification model is shown in Fig. 2. The logical mechanism includes a set of axioms defined based on domains (features and labels), variables, constants (classes), and predicates/model/function (p); details are shown in Table 2, which was inspired by the initial study of Logic Tensor Network [51]. To clarify, the Fig. 2 diagram illustrates the Logic Tensor Network (LTN) based Architecture of the hERG-LTN model, where A) represents several types of feature extraction to input to train the model, B) Illustrates the classification architecture of the LTN model, which has been conceived from official LTN paper, where features $\boldsymbol{x_A}$ and corresponding label $\boldsymbol{l_A}$ (for class A) having as input into MLP and afterwords universal ∀ quantification equation that impacts loss function, then weights update during backpropagation, same true for $\boldsymbol{x_B}$ (all the features related to B class). Finally, SattAgg function computes the decision of prediction as referred C section in Fig 2. Notably, rules/axioms are $\boldsymbol{P(x_A, l_A), P(x_B, l_B)}$ with First Order Logic based constructed in accordance data nature where p refers as model/MLP/predicate. K is denoted as entire knowledge formulated with real/first-order logic. More details can be found on LTN GitHub tutorial section.

In terms of neural network, Multi-Layer Perceptron (MLP) comprised with four layers where input units for CDK is 1024, and Morgan 2048 features, and various types of LLM embeddings features for the individual hERG-LTN training. The hidden layers consist of 16, 8, and 2 units, with Elu activation, a batch size of 64, an Adam optimizer with a learning rate of 0.001, and a seed of 42 with 500 epochs. The LTN knowledge-based setup is detailed in Table 2. In the training phase, the whole dataset was split into 90% for training and 10% for validation and evaluated using completely unseen external test-1 and the hERG-70 benchmark.

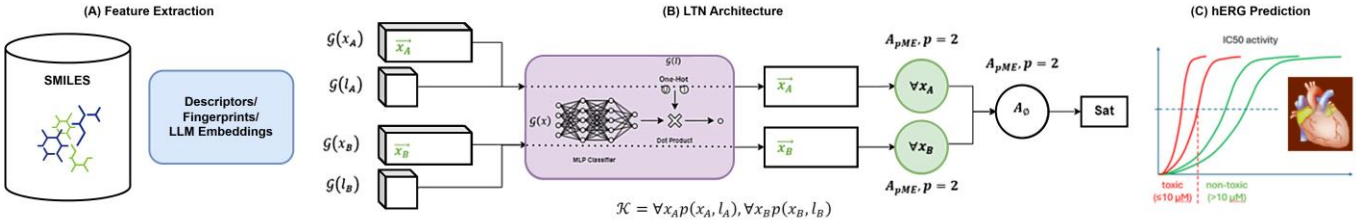

Fig. 2: The diagram illustrates the Logic Tensor Network (LTN) based architecture of the hERG-LTN model [51]

Table 2: LTN Knowledge-based Setting for hERG Classification.

Contents	Blocks
Define Axioms	• $\forall x_A, p(x_A, l_A)$: all the examples of class $A(non-blocks = 0)$ should have a label l_A. • $\forall x_B, p(x_B, l_B)$: all the examples of class B (blocks = 1) should have a label l_B.
Axioms (rules, knowledge base)	$\mathcal{K} = \forall x_A p(x_A, l_A), \forall x_B p(x_B, l_B)$
SatAgg is given by	$\text{SatAgg}_{\phi \in \mathcal{K}} \mathcal{G}_{\boldsymbol{\theta}, x \leftarrow \boldsymbol{D}}(\phi)$
Learning & Loss	$\boldsymbol{L} = \left(1 - \underset{\phi \in \mathcal{K}}{\text{SatAgg}} \mathcal{G}_{\boldsymbol{\theta}, x \leftarrow \boldsymbol{B}}(\phi)\right)$

Here,

The pMeanError aggregator

$$pME(u_1, \ldots, u_n) = 1 - \left(\frac{1}{n}\sum_{i=1}^{n}(1-u_i)^p\right)^{\frac{1}{p}} p \geq 1 \quad (1)$$

(p-mean error) is universal quantification ("for all": ∀): the generalized mean of "the deviations w.r.t (More detail link).

- SatAgg: This stands for "Satisfaction Aggregator" is an operator that aggregates the truth values of the formulas in K (if there are more than one formula).
- ϕ∈K: This part indicates that ϕ (phi) belongs to the set K. ϕ is often used to represent a predicate.
- $\mathcal{G}(\theta)$: This is denoted by grounding ($\mathcal{G}$) with parameters θ. θ represents a set of parameters or weights in a model.
- x←D: D the data set of all examples (domain).
- B is a mini-batch sampled from D.

III. Results

In this section, we analyze the proposed hERG-LTN approach, and its performance compared to existing state-of-the-art models. The main goal is to emphasize the advantages of the hERG-LTN model, which mainly incorporates logical

reasoning for enhanced decision-making in cardiac drug-induced. Tables 3 and 4 present ablation studies across different feature inputs, while Tables 5 and 6 refer to benchmarking comparisons.

Table 3 illustrates the hERG-LTN model's performance based on the External Test-1 dataset. Among the six different types of inputs (CDK, Morgan, MegaMolBART, Gemini-Google, Llama3.2_1.5B-Meta, DeepSeek-QWEN_1.5b), the Morgan features yielded pioneering results, achieving 0.846 (MCC), 0.866 (NPV), 0.931 (ACC), 0.965 (PPV), 0.928 (SPE), 0.933 (SEN), and 0.930 (B-ACC). CDK fingerprints alone are also well competitor with high scores across metrics of 0.721 (MCC), 0.722 (NPV), 0.863 (ACC), 0.961 (PPV), 0.928 (SPE), 0.833 (SEN), and 0.880 (B-ACC). In addition, the experiments incorporated with LLM embeddings (MegaMolBART, Gemini, Llama3.2, DeepSeek-QWEN) demonstrated suboptimal performance compared to Morgan Fingerprint, and LLM based Gemini embedding scored higher than other embedding features.

Table 3: hERG-LTN model performance based on different features input using **UnihERG_DB** training dataset

Evaluation data	Descriptors/ Fingerprint/ Embedding	MCC	NPV	ACC	PPV	SPE	SEN	B-ACC
External Test-1 (pos)	CDK [60]	0.721	0.722	0.863	0.961	0.928	0.833	0.880
	Morgan [64]	**0.846**	**0.866**	**0.931**	**0.965**	**0.928**	**0.933**	**0.930**
	MegaMolBART-Nvidia[52]	0.685	0.684	0.840	0.960	0.928	0.800	0.864
	Gemini-Google [62]	0.700	0.75	0.863	0.928	0.857	0.866	0.861
	Llama3.2_1.5B-Meta [61]	0.660	0.705	0.840	0.925	0.857	0.833	0.845
	DeepSeek-QWEN_1.5b [63]	0.685	0.684	0.840	0.960	0.928	0.800	0.864

Table 4: LTN model performance evaluation based on Arab et al. hERG-70

Evaluation data	Descriptors /Fingerprint/ Embedding	MCC	NPV	ACC	PPV	SPE	SEN	B-ACC
hERG-70	CDK [60]	0.608	0.704	0.791	**0.906**	**0.909**	0.697	0.803
	Morgan [64]	**0.662**	**0.759**	**0.827**	0.899	0.899	**0.777**	**0.833**
	MegaMolBART-Nvidia [52]	0.642	0.748	0.816	0.890	0.880	0.765	0.823
	Gemini-Google [62]	0.534	0.707	0.765	0.822	0.799	0.738	0.768
	Llama3.2_1.5B-Meta [61]	0.535	0.696	0.763	0.836	0.822	0.715	0.769
	DeepSeek-QWEN_1.5b [63]	0.587	0.712	0.786	0.872	0.866	0.723	0.794

In addition, we evaluated model performance with hERG-70 benchmark data presented by Arab et al. [21] (Table 4). The results indicated that the hERG-LTN model performed well with the Morgan features, achieving an MCC of 0.662, NPV of 0.759, ACC of 0.827, SEN of 0.777, and B-ACC of 0.833. MegaMolBART-Nvidia embeddings also showed promising results, with an ACC of 0.816, while CDK was slightly lower. Overall, the results indicate that while all feature sets are practical, Morgan fingerprints provided the most robust performance on this dataset, which is also depicted in Fig. 3 ROC AUC curve.

Table 5: Performance comparison of hERG-LTN classifier on the External Test-1 (pos) dataset.

Evaluation data	Methods	MCC	NPV	ACC	PPV	SPE	SEN	B-ACC
External Test-I (pos) [22]	hERG-LTN	**0.846**	**0.866**	**0.931**	0.965	0.928	**0.933**	**0.930**
	CardioTox [22]	0.599	0.688	0.810	0.893	0.786	0.833	0.810
	M-PNN [23]	0.567	0.656	0.800	0.89	0.786	0.807	0.796
	DeepHIT [24]	0.476	0.643	0.773	0.833	0.643	0.833	0.738
	CardPred [25]	0.193	0.421	0.614	0.760	0.571	0.633	0.602
	OCHEM Predictor-I [28]	0.149	0.333	0.364	**1.000**	**1.000**	0.067	0.534
	OCHEM Predictor-II [28]	0.164	0.351	0.432	0.857	0.929	0.200	0.564
	Pred-hERG 4.2 [29]	0.306	0.538	0.705	0.774	0.500	0.800	0.650

Benchmarking on External Test -1 and hERG-70 dataset

Further, we compare the performance of our hERG-LTN best classification models to the State-of-the-art models as exhibited in the scientific literatures [22]-[25],[28],[29]. We assess with the two different benchmarking datasets, (a) External Test-1 (pos) [22] and (b) hERG-70 [21]. Table 5 shows that the hERG-LTN classifier performed exceptionally well with Morgan features on the External Test-1 (pos), with an 0.846 (MCC), 0.866 (NPV), 0.931 (ACC), 0.965 (PPV), 0.928 (SPE), 0.933 (SEN), and 0.930 (B-ACC). CardioTox was the second-best performer after MCC of 0.599, NPV of 0.688, ACC of 0.810, PPV of 0.893, SPE of 0.786, SEN of 0.833, and B-ACC of 0. 810.OCHEM Predictor-I achieved a PPV of 100% and SPE of 100%.

On the other hand, the hERG-70 benchmark, CardioGenAI (preprint as of Feb 8th) [30], placed the highest ranked, achieving an ACC of 0.835, SEN of 0.862, F1 score of 0.851, and MCC of 0.667, except SPE, and CCR (Table 6). Though, our hERG-LTN classifier yielded the best scores of SPE (0.890) and CCR (0.833), and other metrics closely followed the SOTA models with an ACC of 0.827, SEN of 0.777, SPE of 0.890, F1 score of 0.833, CCR of 0.833, and MCC of 0.662. Alongside, CToxPred-hERG ranked third with an ACC of 0.814, F1 of 0.839, SPE of CCR of 0.807, and MCC of 0.621.

Table 6: Performance comparison of the hERG-LTN classifier on the hERG-70 benchmark dataset.

Evaluation data	Methods	ACC	SEN	SPE	F1	CCR	MCC
hERG-70 [21]	CardioGenAI [30]	**0.835**	0.862	0.803	**0.851**	0.832	**0.667**
	hERG-LTN (ours)	0.827	0.777	**0.89**	0.833	**0.833**	0.662
	CToxPred-hERG [21]	0.814	**0.867**	0.746	0.839	0.807	0.621
	CardioTox [22]	0.812	0.830	0.789	0.831	0.81	0.619
	ADMET1ab 2.0 [26]	0.717	0.716	0.718	0.738	0.717	0.431
	ADMETsar 2.0 [27]	0.685	0.845	0.483	0.750	0.664	0.355
	CardPred [25]	0.561	0.527	0.603	0.570	0.565	0.130

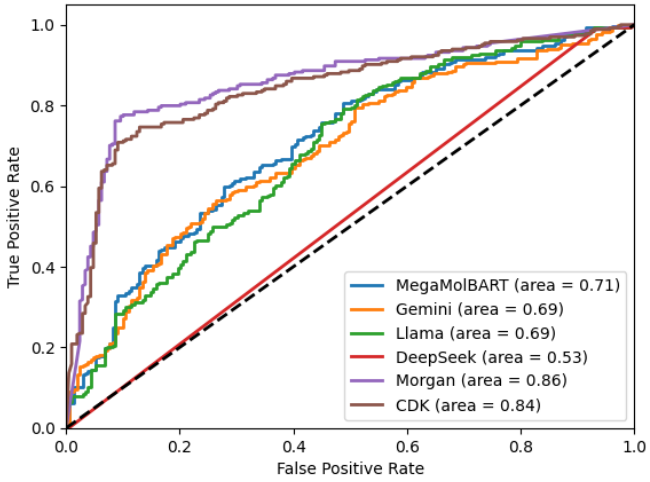

Fig 3: hERG-LTN Model ROC-AUC curve with different language model embeddings.

IV. Discussion

A significant challenge in Cardiotoxicity and AI research is the limited integration of data and knowledge that supports reasoning capabilities. Moreover, current studies fall short of leveraging cutting-edge LLMs for accurate adverse drug prediction. Our findings exhibit that the proposed neuro-symbolic approach, hERG-LTN, trained on the UnihERG_DB dataset, is highly effective in predicting hERG-related Cardiotoxicity. It outperforms existing models such as CardioTox, M-PNN, and DeepHIT on the External Test-1 set (Table 5). Notably, while hERG-LTN ranked just behind CardioGenAI in the hERG-70 benchmark across ACC, F1, and MCC metrics (Table 6), it outperformed all other strategies in terms of Specificity (SPE) and Correct Classification Rate (CCR). Furthermore, the ROC-AUC analysis confirms robust model performance, based on CDK, Morgan, MegaMolBART, Gemini, LLaMA, and DeepSeek-QWEN achieving AUC scores between 0.84 and 0.86 for predicting hERG-70 inhibition.

However, the study does have limitations. Specifically, the hERG-LTN model's slightly lower ACC, F1, and MCC scores on the hERG-70 benchmark that suggests potential areas for further optimization and enhancement.

V. Conclusions

In conclusion, the article emphasizes that the Neuro-Symbolic AI approach called the hERG-LTN classifier establishes a new paradigm for hERG-related Cardiotoxicity evaluation, surpassing existing models. This study explored various methodologies, including molecular fingerprinting,

descriptors, and LLM embeddings. The synthesis of data and knowledge-driven strategies reveals immense potential for propelling precision pharmacogenomics in drug discovery and assessing drug-induced Cardiotoxicity. Future endeavors could involve experimenting with other avant-garde Generative A.I. models, aggregating more heterogeneous datasets, and contemporary Neuro-Symbolic AI frameworks to augment predictive accuracy and model resilience in hERG cardiotoxicity research.

ACKNOWLEDGMENT

We extend our gratitude to SPARC and acknowledge that this project was generously funded through the grants OT2:lOT20D032742-Ol and UMI TR004771.

REFERENCES

[1] M. Shan, C. Jiang, L. Qin, and G. Cheng, "A review of Computational Methods in Predicting HERG channel blockers," ChemistrySelect, vol. 7, no. 31, Aug. 2022, doi: 10.1002/slct.202201221.

[2] P. Mladěnka et al., "Comprehensive review of cardiovascular toxicity of drugs and related agents," Medicinal Research Reviews, vol. 38, no. 4, pp. 1332–1403, Jan. 2018, doi: 10.1002/med.21476.

[3] B. Ahmed, D. Abdelaziz, and B. Abdelmajid, "Advancements in Cardiotoxicity Detection and Assessment through Artificial Intelligence: A Comprehensive Review," IEEE, pp. 1–8, May 2024, doi: 10.1109/iraset60544.2024.10548722.

[4] J. Liu et al., "Machine learning and deep learning approaches for enhanced prediction of hERG blockade: a comprehensive QSAR modeling study," Expert Opinion on Drug Metabolism & Toxicology, vol. 20, no. 7, pp. 665–684, Jul. 2024, doi: 10.1080/17425255.2024.2377593.

[5] I. H. Sanches, S. S. Mendonca, V. M. Alves, R. C. Braga, and C. H. Andrade, "QSAR models for predicting cardiac toxicity of drugs," in Elsevier eBooks, 2023, pp. 351–362. doi: 10.1016/b978-0-443-15339-6.00039-4.

[6] E. Ylipää et al., "hERG-toxicity prediction using traditional machine learning and advanced deep learning techniques," Current Research in Toxicology, vol. 5, p. 100121, Jan. 2023, doi: 10.1016/j.crtox.2023.100121.

[7] P. Delre et al., "Ligand-based prediction of hERG-mediated cardiotoxicity based on the integration of different machine learning techniques," Frontiers in Pharmacology, vol. 13, Sep. 2022, doi: 10.3389/fphar.2022.951083.

[8] Y. Chen, X. Yu, W. Li, Y. Tang, and G. Liu, "In silico prediction of hERG blockers using machine learning and deep learning approaches," Journal of Applied Toxicology, vol. 43, no. 10, pp. 1462–1475, Apr. 2023, doi: 10.1002/jat.4477.

[9] H. Kim, M. Park, I. Lee, and H. Nam, "BayeshERG: a robust, reliable and interpretable deep learning model for predicting hERG channel blockers," Briefings in Bioinformatics, vol. 23, no. 4, May 2022, doi: 10.1093/bib/bbac211.

[10] Y. Zhang et al., "Prediction of hERG K+ channel blockage using deep neural networks," Chemical Biology & Drug Design, vol. 94, no. 5, pp. 1973–1985, Aug. 2019, doi: 10.1111/cbdd.13600.

[11] C. Cai et al., "Deep Learning-Based Prediction of Drug-Induced Cardiotoxicity," Journal of Chemical Information and Modeling, vol. 59, no. 3, pp. 1073–1084, Feb. 2019, doi: 10.1021/acs.jcim.8b00769.

[12] K. M. Saravanan, J.-F. Wan, L. Dai, J. Zhang, J. Z. H. Zhang, and H. Zhang, "A deep learning based multi-model approach for predicting drug-like chemical compound's toxicity," Methods, vol. 226, pp. 164–175, May 2024, doi: 10.1016/j.ymeth.2024.04.020.

[13] Z. Zhu et al., "Two-Dimensional Deep Learning Frameworks for Drug-Induced Cardiotoxicity Detection," ACS Sensors, vol. 9, no. 6, pp. 3316–3326, Jun. 2024, doi: 10.1021/acssensors.4c00654.

[14] J. Song, Y. J. Kim, and C. H. Leem, "Improving the hERG model fitting using a deep learning-based method," Frontiers in Physiology, vol. 14, Feb. 2023, doi: 10.3389/fphys.2023.1111967.

[15] T. M. Creanza, P. Delre, N. Ancona, G. Lentini, M. Saviano, and G. F. Mangiatordi, "Structure-Based Prediction of HERG-Related Cardiotoxicity: A Benchmark study," Journal of Chemical Information and Modeling, vol. 61, no. 9, pp. 4758–4770, Sep. 2021, doi: 10.1021/acs.jcim.1c00744.

[16] I. H. Sanches, R. C. Braga, V. M. Alves, and C. H. Andrade, "Enhancing hERG Risk Assessment with Interpretable Classificatory and Regression Models," Chemical Research in Toxicology, vol. 37, no. 6, pp. 910–922, May 2024, doi: 10.1021/acs.chemrestox.3c00400.

[17] Y. Wang et al., "Capsule networks showed excellent performance in the classification of HERG Blockers/Nonblockers," Frontiers in Pharmacology, vol. 10, Jan. 2020, doi: 10.3389/fphar.2019.01631.

[18] V. B. Siramshetty, D.-T. Nguyen, N. J. Martinez, N. T. Southall, A. Simeonov, and A. V. Zakharov, "Critical assessment of artificial intelligence methods for prediction of HERG channel inhibition in the 'Big Data' Era," Journal of Chemical Information and Modeling, vol. 60, no. 12, pp. 6007–6019, Dec. 2020, doi: 10.1021/acs.jcim.0c00884.

[19] K.-E. Choi, A. Balupuri, and N. S. Kang, "The study on the HERG blocker prediction using chemical fingerprint analysis," Molecules, vol. 25, no. 11, p. 2615, Jun. 2020, doi: 10.3390/molecules25112615.

[20] Y. Yang et al., "Reducing hERG Toxicity Using hERG Classification Model and Fragment-growing Network," ChemRxiv, Mar. 2021, doi: 10.26434/chemrxiv.13153112.v2.

[21] I. Arab, K. Egghe, K. Laukens, K. Chen, K. Barakat, and W. Bittremieux, "Benchmarking of small molecule feature representations for HERG, NAV1.5, and CAV1.2 cardiotoxicity prediction," Journal of Chemical Information and Modeling, vol. 64, no. 7, pp. 2515–2527, Oct. 2023, doi: 10.1021/acs.jcim.3c01301.

[22] A. Karim, M. Lee, T. Balle, and A. Sattar, "CardioTox net: a robust predictor for hERG channel blockade based on deep learning meta-feature ensembles," Journal of Cheminformatics, vol. 13, no. 1, Aug. 2021, doi: 10.1186/s13321-021-00541-z.

[23] M. Shan, C. Jiang, J. Chen, L.-P. Qin, J.-J. Qin, and G. Cheng, "Predicting hERG channel blockers with directed message passing neural networks," RSC Advances, vol. 12, no. 6, pp. 3423–3430, Jan. 2022, doi: 10.1039/d1ra07956e.

[24] J. Y. Ryu, M. Y. Lee, J. H. Lee, B. H. Lee, and K.-S. Oh, "DeepHIT: a deep learning framework for prediction of hERG-induced cardiotoxicity," Bioinformatics, vol. 36, no. 10, pp. 3049–3055, Jan. 2020, doi: 10.1093/bioinformatics/btaa075.

[25] H.-M. Lee et al., "Computational determination of hERG-related cardiotoxicity of drug candidates," BMC Bioinformatics, vol. 20, no. S10, May 2019, doi: 10.1186/s12859-019-2814-5.

[26] G. Xiong et al., "ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties," Nucleic Acids Research, vol. 49, no. W1, pp. W5–W14, Mar. 2021, doi: 10.1093/nar/gkab255.

[27] H. Yang et al., "admetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties," Bioinformatics, vol. 35, no. 6, pp. 1067–1069, Aug. 2018, doi: 10.1093/bioinformatics/bty707.

[28] X. Li, Y. Zhang, H. Li, and Y. Zhao, "Modeling of the HERG K+ channel blockage using Online Chemical Database and Modeling Environment (OCHEM)," Molecular Informatics, vol. 36, no. 12, Aug. 2017, doi: 10.1002/minf.201700074.

[29] R. C. Braga et al., "Pred-HERG: a novel web-Accessible computational tool for predicting cardiac toxicity," Molecular Informatics, vol. 34, no. 10, pp. 698–701, Jul. 2015, doi: 10.1002/minf.201500040.

[30] G. W. Kyro, M. T. Martin, E. D. Watt, and V. S. Batista, "CardioGenAI: A Machine Learning-Based Framework for Re-Engineering Drugs for Reduced HERG Liability," Research Square (Research Square), Sep. 2024, doi: 10.21203/rs.3.rs-4896795/v1.

[31] K. Hamilton, A. Nayak, B. Božić, and L. Longo, "Is neuro-symbolic AI meeting its promises in natural language processing? A structured review," Semantic Web, vol. 15, no. 4, pp. 1265–1306, Nov. 2022, doi: 10.3233/sw-223228.

[32] G. G. Towell and J. W. Shavlik, "Knowledge-based artificial neural networks," Artificial Intelligence, vol. 70, no. 1–2, pp. 119–165, Oct. 1994, doi: 10.1016/0004-3702(94)90105-8.

[33] Jang, S., Girard, M.J., & Thiéry, A.H. (2022). Explainable and Interpretable Diabetic Retinopathy Classification Based on Neural-Symbolic Learning. ArXiv, abs/2204.00624.

[34] Yang, F., Yang, Z., & Cohen, W.W. (2017). Differentiable Learning of Logical Rules for Knowledge Base Reasoning. Neural Information Processing Systems.

[35] P. Hohenecker and T. Lukasiewicz, "Ontology Reasoning with Deep Neural Networks," Journal of Artificial Intelligence Research, vol. 68, Jul. 2020, doi: 10.1613/jair.1.11661.

[36] L. Luo, G. Zhang, H. Xu, Y. Yang, C. Fang, and Q. Li, "INSIGHT: End-to-End Neuro-Symbolic Visual Reinforcement Learning with Language Explanations," arXiv (Cornell University). https://arxiv.org/abs/2403.12451

[37] P. Kora, K. Meenakshi, K. Swaraja, A. Rajani, and M. K. Islam, "Detection of Cardiac arrhythmia using fuzzy logic," Informatics in Medicine Unlocked, vol. 17, p. 100257, Jan. 2019, doi: 10.1016/j.imu.2019.100257.

[38] R. Maclin and J. W. Shavlik, "Using Knowledge-Based Neural Networks to Improve Algorithms: Refining the Chou–Fasman Algorithm for Protein Folding," Springer, vol. 11, no. 2/3, pp. 195–215, Jan. 1993, doi: 10.1023/a:1022609403428.

[39] J. Wang, J. Zhang, Y. Cai, and L. Deng, "DEEPMIR2GO: Inferring functions of human MicroRNAs using a deep Multi-Label Classification model," International Journal of Molecular Sciences, vol. 20, no. 23, p. 6046, Nov. 2019, doi: 10.3390/ijms20236046.

[40] K. Yi, J. Wu, C. Gan, A. Torralba, P. Kohli, and J. Tenenbaum, "Neural-Symbolic VQA: Disentangling Reasoning from Vision and Language Understanding," 2018. https://api.semanticscholar.org/CorpusID:52919654

[41] S. Amizadeh, H. Palangi, O. Polozov, Y. Huang, and K. Koishida, "Neuro-Symbolic Visual Reasoning: Disentangling 'Visual' from 'Reasoning,'" arXiv (Cornell University), Jan. 2020, doi: 10.48550/arxiv.2006.11524.

[42] R. Riegel et al., "Logical neural networks," arXiv (Cornell University), Jan. 2020, doi: 10.48550/arxiv.2006.13155.

[43] G. Towell and J. W. Shavlik, "Interpretation of Artificial Neural Networks: Mapping Knowledge-Based Neural Networks into Rules," Neural Information Processing Systems, vol. 4, pp. 977–984, Dec. 1991, [Online]. Available: http://papers.nips.cc/paper/546-interpretation-of-artificial-neural-networks-mapping-knowledge-based-neural-networks-into-rules.pdf

[44] A. Lavin, "Neuro-Symbolic neurodegenerative disease modeling as probabilistic programmed deep kernels," in Studies in computational intelligence, 2022, pp. 49–64. doi: 10.1007/978-3-030-93080-6_5.

[45] K. Dobosz and W. Duch, "Fuzzy symbolic dynamics for neurodynamical systems," in Lecture notes in computer science, 2008, pp. 471–478. doi: 10.1007/978-3-540-87559-8_49.

[46] F. Arabshahi, J. Lee, M. Gawarecki, K. Mazaitis, A. Azaria, and T. Mitchell, "Conversational Neuro-Symbolic commonsense reasoning," arXiv (Cornell University), Jan. 2020, doi: 10.48550/arxiv.2006.10022.

[47] Z. Yang, A. Ishay, and J. Lee , "NeurASP: Embracing Neural Networks into Answer Set Programming," ACM, Art. no. 243, Jan. 2021, doi: 10.5555/3491440.

[48] S. Jiaxin, Z. Hanwang, and L. Juanzi, "Explainable and Explicit Visual Reasoning over Scene Graphs," arXiv (Cornell University), Jan. 2018, doi: 10.48550/arxiv.1812.01855.

[49] J. Xu, Z. Zhang, T. Friedman, Y. Liang, and G. Van Den Broeck, "A Semantic Loss Function for Deep Learning with Symbolic Knowledge," arXiv (Cornell University), Jan. 2017, doi: 10.48550/arxiv.1711.11157.

[50] J. Mao, C. Gan, P. Kohli, J. B. Tenenbaum, and J. Wu, "The Neuro-Symbolic Concept Learner: Interpreting scenes, words, and sentences from natural supervision," arXiv (Cornell University), Jan. 2019, doi: 10.48550/arxiv.1904.12584.

[51] S. Badreddine, A. D. Garcez, L. Serafini, and M. Spranger, "Logic Tensor networks," Artificial Intelligence, vol. 303, p. 103649, Dec. 2021, doi: 10.1016/j.artint.2021.103649.

[52] "MegaMolBART," NVIDIA BioNeMo Framework. https://docs.nvidia.com/bionemo-framework/1.10/models/megamolbart.html

[53] B. Zdrazil et al., "The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods," Nucleic Acids Research, vol. 52, no. D1, pp. D1180–D1192, Nov. 2023, doi: 10.1093/nar/gkad1004.

[54] M. K. Gilson, T. Liu, M. Baitaluk, G. Nicola, L. Hwang, and J. Chong, "BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology," Nucleic Acids Research, vol. 44, no. D1, pp. D1045–D1053, Oct. 2015, doi: 10.1093/nar/gkv1072.

[55] S. Kim et al., "PubChem 2023 update," Nucleic Acids Research, vol. 51, no. D1, pp. D1373–D1380, Oct. 2022, doi: 10.1093/nar/gkac956.

[56] S. D. Harding et al., "The IUPHAR/BPS Guide to PHARMACOLOGY in 2024," Nucleic Acids Research, vol. 52, no. D1, pp. D1438–D1449, Oct. 2023, doi: 10.1093/nar/gkad944.

[57] S. Wang, H. Sun, H. Liu, D. Li, Y. Li, and T. Hou, "ADMET Evaluation in Drug Discovery. 16. Predicting HERG blockers by combining multiple pharmacophores and machine learning approaches," Molecular Pharmaceutics, vol. 13, no. 8, pp. 2855–2866, Jul. 2016, doi: 10.1021/acs.molpharmaceut.6b00471.

[58] D. Hossain and J. Y. Chen, "A study on Neuro-Symbolic Artificial Intelligence: Healthcare Perspectives," arXiv.org, Mar. 23, 2025. https://arxiv.org/abs/2503.18213

[59] D. Hossain and J. Y. Chen, "NESYDPP4-QSAR: A Neuro-Symbolic AI approach for potent DPP-4-Inhibitor discovery in diabetes treatment," bioRxiv (Cold Spring Harbor Laboratory), Apr. 2025, doi: 10.1101/2025.03.31.646336.

[60] C. W. Yap, "PaDEL-descriptor: An open-source software to calculate molecular descriptors and fingerprints," Journal of Computational Chemistry, vol. 32, no. 7, pp. 1466–1474, Dec. 2010, doi: 10.1002/jcc.2170.

[61] M. Ettaleb, M. Kamel, V. Moriceau, and N. Aussenac-Gilles, "The Llama 3 Herd of Models," Arxiv, Nov. 2024, doi: 10.48550/arxiv.2407.21783.

[62] G. Team et al., "Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context," arXiv.org, Mar. 08, 2024. https://arxiv.org/abs/2403.05530

[63] DeepSeek-Ai et al., "DeepSeek-V3 Technical Report," arXiv.org, Dec. 27, 2024. https://arxiv.org/abs/2412.19437

[64] H. L. Morgan, "The generation of a Unique Machine Description for Chemical Structures-A technique developed at Chemical Abstracts Service.," Journal of Chemical Documentation, vol. 5, no. 2, pp. 107–113, May 1965, doi: 10.1021/c160017a018.