Received: 13 November 2023 | Revised: 21 December 2023 | Accepted: 29 December 2023 | Published online: 26 January 2024

RESEARCH ARTICLE

Medinformatics
2024, Vol. 1(3) 103–111
DOI: 10.47852/bonviewMEDIN42022049

BON VIEW PUBLISHING

CLOP-hERG: The Contrastive Learning Optimized Pre-trained Model for Representation Learning in Predicting Drug-Induced hERG Channel Blockers

Shida He[1], Xiucai Ye[1,*] and Tetsuya Sakurai[1]

[1]*Department of Computer Science, University of Tsukuba, Japan*

Abstract: During drug development, ensuring that drug molecules do not block the hERG (human Ether-à-go-go-Related Gene) channel is critical. If this channel is blocked, it can cause many cardiovascular-related diseases. However, traditional experimental detection methods are expensive and time-consuming. In this work, we proposed a novel deep learning framework CLOP-hERG that combines contrastive learning with the RoBERTa pre-trained model to predict whether drug molecules will block the hERG channel. We employed data augmentation techniques on molecular structures to ensure that our model can capture the multifaceted information of the molecules. Besides, we used a contrastive learning strategy to enable the model to learn meaningful molecular features from large unlabeled datasets. The RoBERTa pre-trained model played a pivotal role in this process, giving our model with a robust representational learning capability. The model, obtained through contrastive learning, was further fine-tuned to enhance high-precision prediction of hERG channel blockers. Through a series of experiments, we demonstrated the effectiveness of CLOP-hERG. This work provides a novel and an effective strategy for predicting hERG blockers and provides some insights for other similar pharmaceutical tasks.

Keywords: hERG, contrastive learning, RoBERTa, representation learning

1. Introduction

Cardiovascular diseases have persistently posed a profound challenge to global health, profoundly undermining human vitality and quality of life. The hERG gene [1], encoding the α subunit Kv11.1 protein of potassium ion channels, is integral to cardiac electrical activity. Certain pharmaceutical compounds are known to inhibit the hERG channel, precipitating severe cardiac complications, such as prolonged QT intervals, sudden cardiac arrest, and arrhythmias, thus presenting a considerable obstacle in drug development. The adverse effects on the hERG channel have necessitated the withdrawal of numerous drugs from the market, underscoring the imperative for rigorous toxicity assessments in the drug discovery pipeline. For instance, medicines such as astemizole, terfenadine, and cisapride have been rendered obsolete due to their propensity to block the hERG channel [2–4], detrimentally impacting public health and inflicting substantial economic losses on the pharmaceutical industry. Consequently, evaluating the potential of compounds to block the hERG channel has crystallized as a pivotal research focus in the advancement of new pharmaceuticals.

In recent years, with the advancement of technology, there have been great breakthroughs in many disciplines such as biology, chemistry, and computers science [5–7]. These advancements have spurred the development of innovative methods for predicting drug interactions and the hERG channel. These computational techniques enable the rapid assessment of potential drug risks [8, 9], allowing researchers to gain a deeper understanding of the complex relationship between molecular structure and biological activity. However, challenges related to data quality, model interpretability, and prediction accuracy still exist. In drug discovery research, most methods rely on molecular fingerprint [10] and related approaches. The fundamental principle of molecular fingerprint involves encoding the chemical structure of a molecule into a fixed-length vector, facilitating computational processing and analysis. These vectors can be viewed as the "identity" of the molecule, with each individual (or element) representing certain specific structural features or properties of the molecule. Molecular fingerprints are widely used in cheminformatics, especially in the field of drug discovery, for tasks such as molecular similarity comparison, virtual screening, and compound classification. Compared to traditional molecular fingerprinting, deep learning [7, 11, 12] represents a paradigm shift; autonomously learning complex, hierarchical feature representations from data, and quantitative simulations can avoid the problem of intrinsic information loss in manual feature engineering. It also allows for extracting complex information from molecules, further refining the prediction model and making the drug discovery process more accurate and efficient.

*Corresponding author: Xiucai Ye, Department of Computer Science, University of Tsukuba, Japan. Email: yexiucai@cs.tsukuba.ac.jp

On the other hand, traditional methods such as patch-clamp electrophysiology have been extensively utilized to assess the hERG channel blocking characteristics of candidate drugs. However, these procedures are costly and time-consuming when screening all potential drugs individually. Recently, the advent of machine learning, particularly deep learning techniques, has been applied across various bioinformatics domains, such as protein structure prediction, gene interaction profiling, and molecular generation. Ogura et al. [13] used a support vector machine with descriptor selection based on Non-dominated Sorting Genetic Algorithm-II to optimize the descriptor set for maximum prediction performance with the minimal number of descriptors and then used two selected descriptors and ECFP4 fingerprints to predict molecular blockers. Wang et al. [14] integrate multiple molecular fingerprint features and utilize a multi-head attention mechanism for comprehensive graph feature extraction, ultimately classifying compounds as hERG blockers or non-blockers using a fully connected neural network. Karim et al. [15] proposed a deep learning framework based on step-wise training to predict small molecules' hERG channel blocking activity. Their approach utilizes five individual deep learning base models with their respective base features and a separate neural network to combine the outputs of the five base models.

In this study, we proposed a feature representation method specifically for hERG molecules by adopting a contrastive learning optimization strategy based on a pre-trained model [16–18] RoBERTa [10] from PubChem [19]. Our approach draws upon the method of the SimCLR [20] network for contrastive learning. This enables us to extract more distinctive and representative features from the pre-trained model. We further refined the model through fine-tuning, ensuring that these representations are both efficient and accurate for the task of determining whether a molecule is a hERG channel blocker. In contrast to other researchers who predominantly utilize feature fusion methods for identifying molecular hERG channel blockers, our CLOP-hERG method is designed to extract molecular features. Moreover, our method combines contrastive learning with the pre-trained model to achieve a deeper and more precise feature extraction. In addition, users can use the friendly Python script we provide to obtain molecular representations. We believe this innovative approach will provide fresh insights into the study of hERG and other drug molecules.

2. Materials and Methods

2.1. Datasets

In this work, we initially employed the Therapeutics Data Commons database [21], a repository offering an array of drug molecule data for the evaluation of AI methodologies. To enhance the fitting capacity of our model, we gathered additional data from the hERG dataset constructed by Kim et al. [22], comprising information sourced from databases such as ChEMBL, PubChem, and BindingDB. Together, these datasets constitute a substantial collection featuring 26,006 unique drug molecules. To ensure the accuracy of our analysis and the high quality of the dataset, we commenced with a data cleansing step. During the preprocessing phase, we meticulously purged duplicate entries from the dataset to guarantee the unique presence of each molecular structure. Subsequently, we employed the Tanimoto similarity coefficient (Threshold 0.85) as a criterion to filter the degree of similarity between molecules within the dataset. As a prevalent metric for assessing molecular similarity, the Tanimoto similarity quantifies the resemblance between two molecules by calculating the overlap between their fingerprints. We then eliminated molecules with high similarity using RDKit's morgan fingerprint [23] to prevent biases in model training and validation that could arise from substantial compound homogeneity. Through data cleansing and filtering steps, we ultimately compiled a dataset comprising 5,736 positive samples and 5,305 negative samples.

2.2. Methods

2.2.1. The network architecture of the proposed CLOP-hERG

This work introduces a novel contrastive learning framework, CLOP-hERG, specifically architected to address the challenge of molecular toxicity prediction with enhanced efficiency. The CLOP-hERG architecture integrates several components: a data augmentation module, a pre-trained natural language processing model (RoBERTa), a contrastive learning mechanism, and a subsequent multilayer perceptron (MLP) classifier. The data augmentation module employs random masking and sequence shuffling techniques to enrich the training sample set, bolstering the model's generalization capabilities. The pre-trained RoBERTa model can harness to extract robust feature representations, which are refined further through a contrastive learning mechanism—assuring the model can discern differences between molecules. Finally, the *t*-distributed stochastic neighbor embedding (*t*-SNE) method is applied to visualize the extracted features, providing an intuitive assessment of the model's performance. Ultimately, these features are utilized to train a MLP classifier, enabling precise predictions. The architecture of our model CLOP-hERG is shown in Figure 1.

2.2.2. Data augmentation

In this study, a data augmentation strategy was employed to enrich the molecular dataset, thereby enhancing the model's predictive capabilities regarding hERG channel blockers. This augmentation occurs in a two-step process: initially, we introduce stochasticity into the molecular representation by masking random atoms with a "MASK" token, thereby mimicking potential variations in interaction sites. This is followed by shuffling the order of bits to increase further the structural diversity represented in the dataset. To construct a training set compatible with contrastive learning techniques, we define our positive sample pairs as two distinct augmented instances derived from the same original molecule—essentially, the same molecule subjected to data augmentation twice. Negative sample pairs, on the other hand, are formed by applying the augmentation process to two different molecules, resulting in variations that reflect their inherent structural disparities. This data augmentation approach is pivotal for expanding the dataset and providing a robust framework for the model to discern and learn from the differences between molecular interactions.

2.2.3. Contrastive learning

In the contrastive learning method of our investigation, we were inspired by the method of SimCLR, implementing a self-supervised learning strategy designed to extract meaningful representations from data. This model is trained to discern and differentiate between various augmented samples. During this process, the model learns by amplifying the concordance of positive sample pairs—distinct augmentations originating from an identical molecule—while concurrently diminishing the similarity among negative sample pairs derived from disparate molecules. Employing a contrastive loss function compels the model to cluster analogous samples in the feature space and segregate

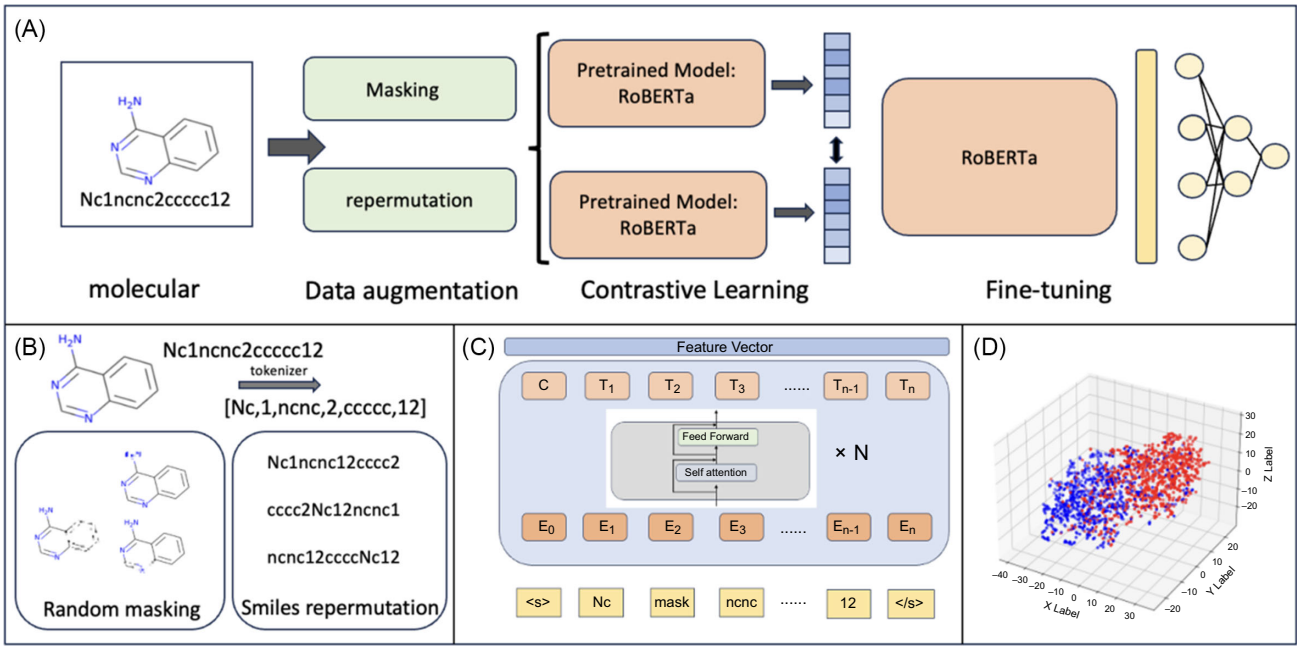

Figure 1. The flowchart describes the overall implementation of the proposed contrast learning approach for predicting hEFG channel blockers. (A) The architecture of proposed CLOP-hERG. (B) Data augmentation strategies in CLOP-hERG. (C) The architecture of RoBERTa. (D) The result of representations visualization.

dissimilar ones, thereby getting better feature representations that can capture the distinguishing attributes of molecules.

2.2.4. Pre-trained model

Pre-trained models are well-used by researchers for their ability to capture rich information from vast datasets. By being trained on large-scale chemical datasets, these models can learn extensive chemical language and structural features, thereby acquiring powerful feature extraction capabilities. The pre-trained RoBERTa (from Hugging Face: PubChem10M_SMILES_BPE_450k) is a deep learning model specifically designed for molecular sequence analysis. Built upon the robust RoBERTa framework, it has been intensively trained on a substantial dataset from the public chemical database: PubChem. This database encompasses various chemical structures and properties, offering a rich tapestry of molecular information. Utilizing byte pair encoding with many subword units derived from the SMILES notation, the model can recognize and interpret a comprehensive chemical vocabulary. This is important in capturing the interactions of molecules, which is essential for accurately predicting molecular characteristics, bioactivities, and potential toxicities. The pre-trained RoBERTa model was trained on a big chemical dataset, which significantly bolsters its performance and dependability for subsequent applications.

2.2.5. Fine-tuning

We trained a MLP classifier using the extracted representations. The MLP is a simple yet powerful, fully connected neural network capable of capturing complex nonlinear patterns through its hidden layers. Within our framework, the MLP receives feature vectors obtained from the RoBERTa and contrastive learning processes as input and learns how to effectively classify based on these features to predict the toxicity of molecules. With the MLP classifier, we achieved high-accuracy predictions ability, demonstrating the potential of the entire framework for practical application.

2.2.6. Loss function

In this work, our comparative learning refers to the SimCLR network. SimCLR employed is the Normalized Temperature-scaled Cross-Entropy Loss (NT-Xent loss). In essence, this loss function encourages the model to align the feature vectors of positive pairs more closely in the latent space while concurrently dispersing the feature vectors of negative pairs. Minimizing this loss drives the model to learn a feature representation wherein the vectors of positive pairs are closer in the high-dimensional space than those of negative pairs, thereby facilitating the model's ability to discern between molecular structures at the feature level. This loss function is articulated as follows:

$$l_{i,j} = -\log \frac{\exp\left(\mathrm{sim}\left(\mathbf{z_i}, \mathbf{z_j}\right)/\tau\right)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} \exp\left(\mathrm{sim}\left(\mathbf{z_i}, \mathbf{z_k}\right)/\tau\right)}$$

where z_i and z_j represent the feature vectors of two different augmentations of the same input sample, transformed through a neural network. The function *sim* (cosine similarity) quantifies the similarity between z_i and z_j. The variable τ denotes a temperature scaling parameter, affecting the separation of similarity scores. The denominator aggregates the exponential similarities of all negative pairs within the batch, wherein $\mathbb{1}_{[k \neq i]}$ serves as an indicator function, equating to 1 when k ≠ i and ensuring that positive pairs are not inadvertently treated as their own negatives.

Cross-entropy loss is commonly employed during the fine-tuning phase of a model, particularly in supervised learning tasks such as molecular toxicity prediction. This loss function measures the discrepancy between the predicted probability distribution (outputted by the model) and the true distribution, where the true distribution is the distribution of the actual labels. The Cross-entropy loss for a binary classification problem can be formulated as:

$$L(y, \hat{y}) = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

where y is the actual label (0 or 1) of the sample. $\hat{y}$ is the predicted probability that the sample belongs to class 1, as outputted by the model.

Evaluate metrics. To evaluate the performance of our proposed predictor, we used traditional evaluation metrics commonly used in binary classification tasks, including accuracy (ACC), $F1$ score, precision, recall, Matthews correlation coefficient (MCC), and area under the receiver operating characteristic curve (AUC). The metrics are calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FN)(TP + FP)(TN + FP)(TN + FN)}}$$

where TP (true positive) and TN (true negative) represent the numbers of correctly predicted positive samples and negative samples, respectively; FP (false positive) and FN (false negative) represent the numbers of incorrectly predicted positive samples and negative samples, respectively. In addition, AUC is also used as one of the evaluation indicators of this model.

3. Results and Discussion

3.1. Performance evaluation

This section explored the effectiveness of our method by comparing the performance with other methods in predicting drug activity. Table 1 provides a comparative analysis of various computational models tested for their ability to predict hERG blockers, a crucial aspect in the evaluation of potential drug candidates. Each model is assessed across multiple performance metrics, namely accuracy (ACC), $F1$ score, recall, precision, AUC, and MCC.

The CLOP-hERG model stands out with superior performance across all metrics, achieving an accuracy of 0.804, an $F1$ score of 0.803, a recall of 0.803, and precision of 0.803, indicating a high level of reliability and balance between sensitivity and specificity. Its AUC score is 0.878, suggesting excellent ability to distinguish hERG blockers. The MCC value, a balanced measure that considers all four categories of the confusion matrix, is 0.608, reflecting a strong correlation between the model's predictions and the actual data. In comparison, other models show varied performance. For example, in the comparative approach, TextCNN shows slightly better results with an $F1$ score of 0.779 and an AUC of 0.842, but, overall, it is not as good as the results of CLOP-hERG. Overall, the CLOP-hERG model's outperformance suggests its robustness and potential as a reliable tool in drug safety assessment, specifically for evaluating hERG-related blockers.

3.2. Ablation experiment

In our ablation studies, the CLOP-hERG model demonstrated superior performance, significantly outperforming other variants. The results are shown in Figure 2, and the pre-trained RoBERTa model has established a strong baseline, but the addition of contrastive learning without the support of the pre-trained model did not bring the expected gain. The separate RoBERTa model confirms the key role of pre-training in improving model performance. The CLOP-hERG model integrating all components has achieved obvious advantages in various indicators, which shows the superiority of this method in prediction accuracy and robustness.

To further substantiate the performance of our method, we analyzed the results of our ablation study with five repetitions using the statistics methods [33, 34]. These tests compared the performance metrics of different components of CLOP-hERG, and the p-values obtained from the t-tests were visualized in a heatmap. The final result was significantly lower than the 0.05 threshold, indicating robust and significant performance differences and highlighting the superiority of our CLOP-hERG method in capturing the molecular characterizations of hERG blockers. The results are shown in Figure 3.

3.3. Comparison with other feature extraction methods

In our experiments, CLOP-hERG outperformed a range of commonly used feature extraction methods. As is shown in Figure 4, it achieved the highest scores in accuracy (ACC), $F1$ score, recall, precision, MCC, and AUC. The high accuracy (ACC) of 0.804 achieved by CLOP-hERG underscores its exceptional capability in classifying hERG blockers. This result reflects the reliability of CLOP-hERG in handling molecular data. In contrast, methods like OneHotFeaturizer that utilize simplistic encoding approaches fall short in capturing the complexity of molecular data, leading to lower performance metrics. For its performance on MCC, a comprehensive

Table 1. Performance comparison with other model methods

Model	ACC	$F1$	Recall	Precision	AUC	MCC
LR [24]	0.748	0.749	0.746	0.751	0.809	0.496
SVM [25]	0.754	0.756	0.753	0.759	0.829	0.509
TextCNN [26]	0.775	0.779	0.745	0.790	0.842	0.551
Bi-LSTM [27]	0.751	0.756	0.760	0.751	0.818	0.501
BERT [28]	0.733	0.747	0.731	0.762	0.789	0.464
GCN [29]	0.734	0.734	0.734	0.734	0.810	0.468
GAT [30]	0.718	0.718	0.718	0.718	0.780	0.436
DMFGAM [14]	0.768	0.758	0.730	0.789	0.845	0.537
hERG-Attn [31]	0.723	0.728	0.747	0.711	0.792	0.446
Smiles2vec [32]	0.745	0.731	0.806	0.777	0.819	0.492
CLOP-hERG	**0.804**	**0.803**	**0.803**	**0.803**	**0.878**	**0.608**

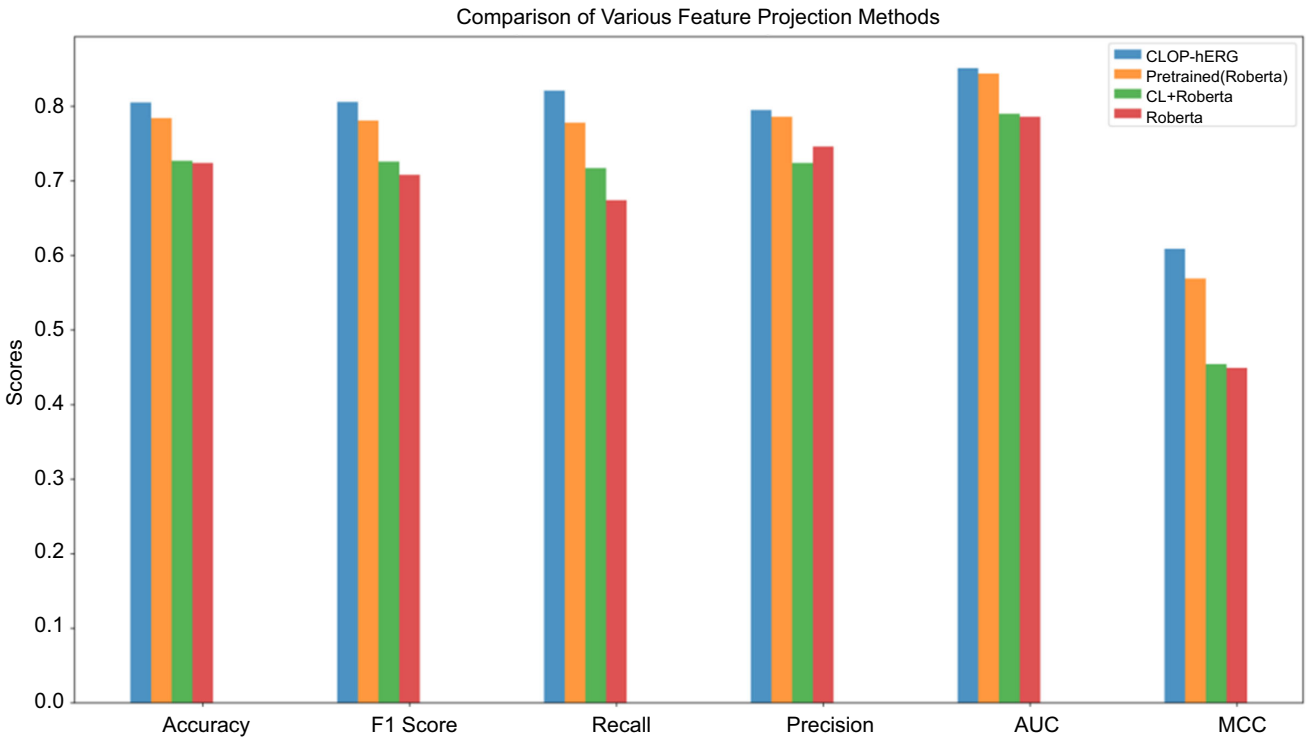

Figure 2. Ablation experiment evaluation

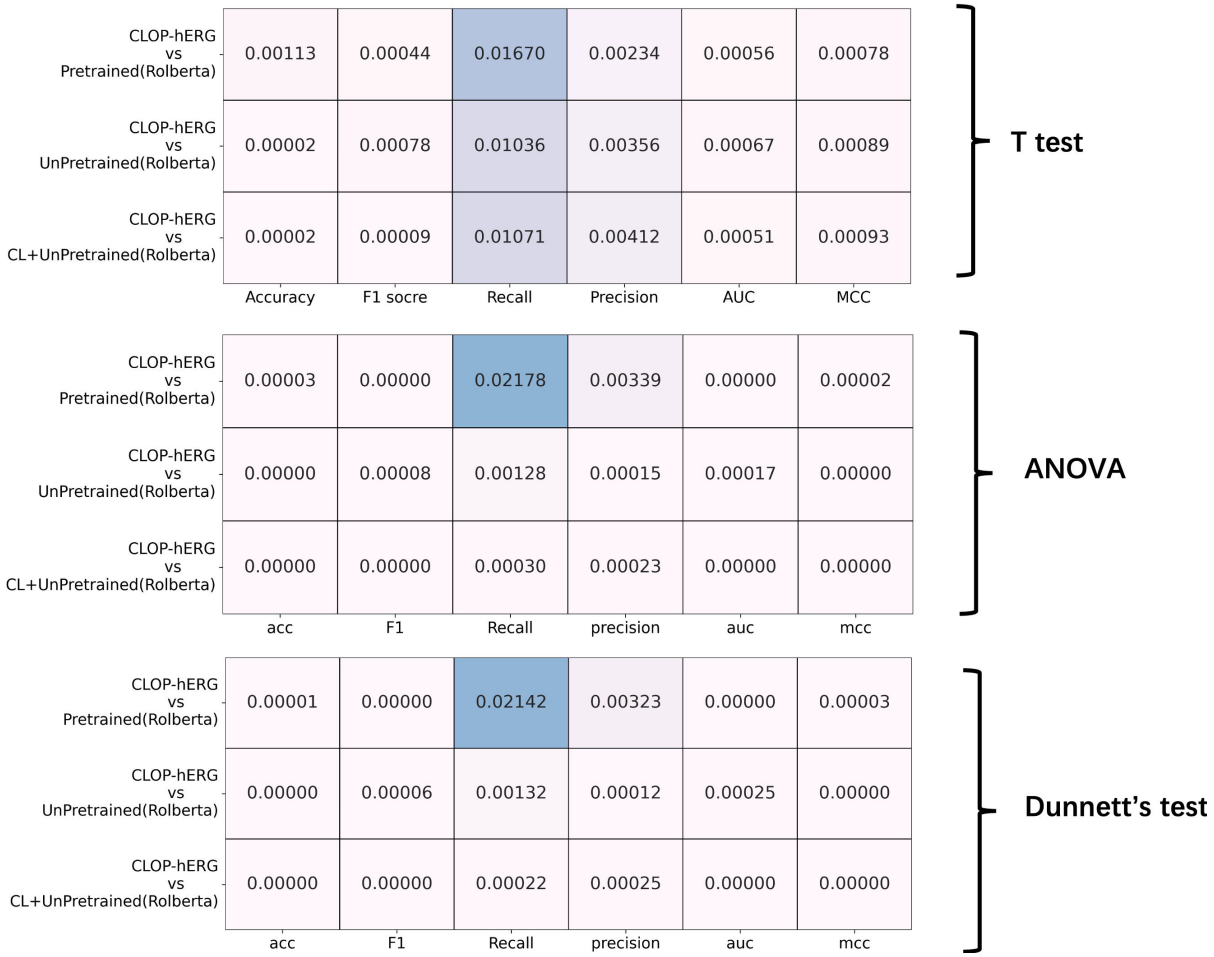

Figure 3. Heatmap of *p*-values for the ablation experiment (repeat 5 times)

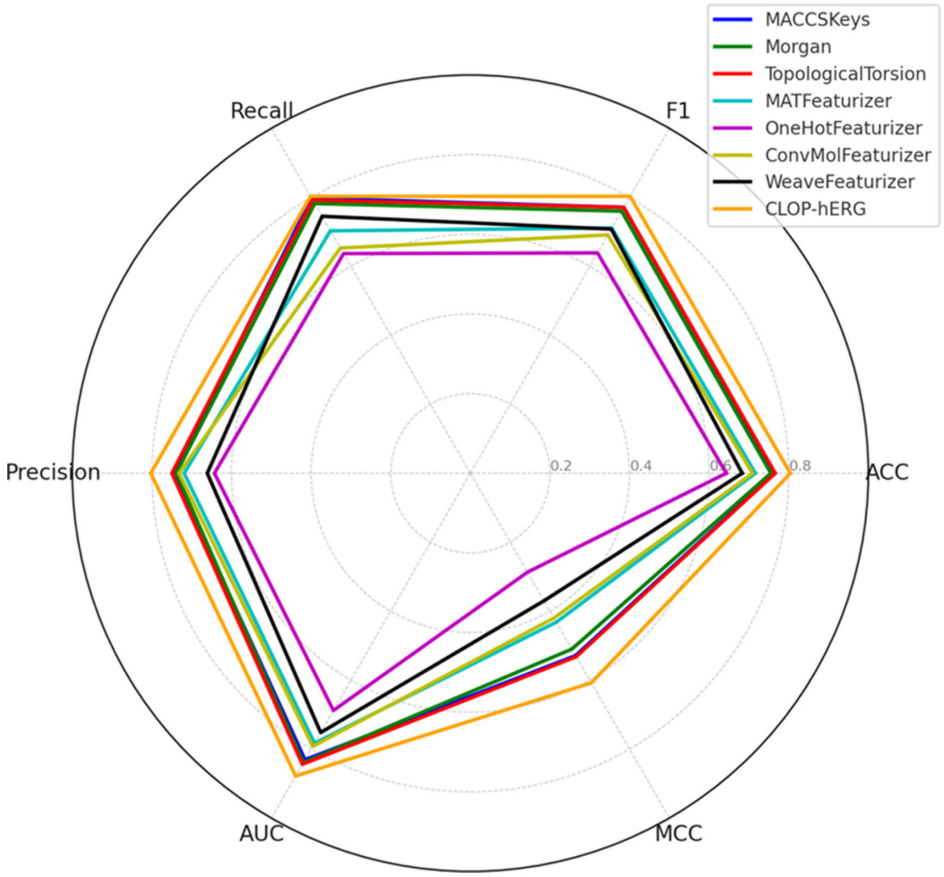

Figure 4. Performance comparison with other feature extraction methods (MLP)

indicator of classification performance, CLOP-hERG reached 0.608, which is much higher than other methods. These results demonstrate that CLOP-hERG can not only effectively capture the characteristics of molecules but also demonstrate its robustness and accuracy on different evaluation metrics. The diversity of methods compared, ranging from fingerprint-based approaches like MACCSKeys to more complex featurizers like ConvMolFeaturizer, highlights the adaptability of CLOP-hERG. It demonstrates not only its efficacy over traditional methods but also over deep learning-based approaches that are represented by ConvMolFeaturizer and WeaveFeaturizer. This indicates that CLOP-hERG's approach to feature representation and classification effectively captures the nuanced structural and chemical properties of molecules that are vital for accurate hERG blockers prediction. These comparison methods are available from RDKit [23] and Deepchem [35]. In the future work, we will explore integrating the model CLOP-hERG with other classification models and feature extraction methods to examine whether a feature fusion method can further enhance performance.

3.4. Feature analysis by SHAP

In this work, we employed shapley additive explanations (SHAP) values [36] to elucidate the underpinnings of our predictive model's decision-making process [37]. This analytical approach permitted us to dissect and quantify the individual contributions of features to specific predictions, thereby unveiling the elements that exert the most pronounced influence on the model's outputs. Figure 5 delineates several features that impart substantial positive impetus to the predictive accuracy of the model for positive samples. Manifested as red bars within the figure, the magnitude and value of these features are directly proportional to their positive impact on model predictions. Notably, the feature "NCCOc" registers a significant positive SHAP value, signifying a potent positive contribution to the model's predictive prowess. Concomitantly, attributes such as "CCCCCCCN" and "CCCOCCn" also emerge as substantial positive contributors. In contrast, our analysis concurrently identified a subset of features that bear negative SHAP values, such as "cncs" and "SCC", indicative of their detrimental influence on the model's forecasting accuracy. By getting insights derived from the analysis of molecular features, or exploring these molecular fragments for potential biomarkers, researchers can design and screen candidate drugs more efficiently, thus accelerating the decision-making process in the early stages of drug development.

3.5. Representation visualization

Figure 6 shows a 3D visualization of COLP-hERG model features using *t*-SNE [38]. For the representations we extracted from the RoBERTa model before fine-tuning, we further employed the *t*-SNE technique for dimensionality reduction and visualization. *t*-SNE is a potent nonlinear dimensionality reduction method that can transform high-dimensional data into points within a low-dimensional space, with the distribution of these points reflecting the structure inherent in the original dataset. The left picture is the feature distribution of the training dataset, and the right picture is the feature distribution of the test dataset. The blue and red points represent different categories. They are clustered more closely, indicating that the model has learned the features that distinguish these two types of data during training.

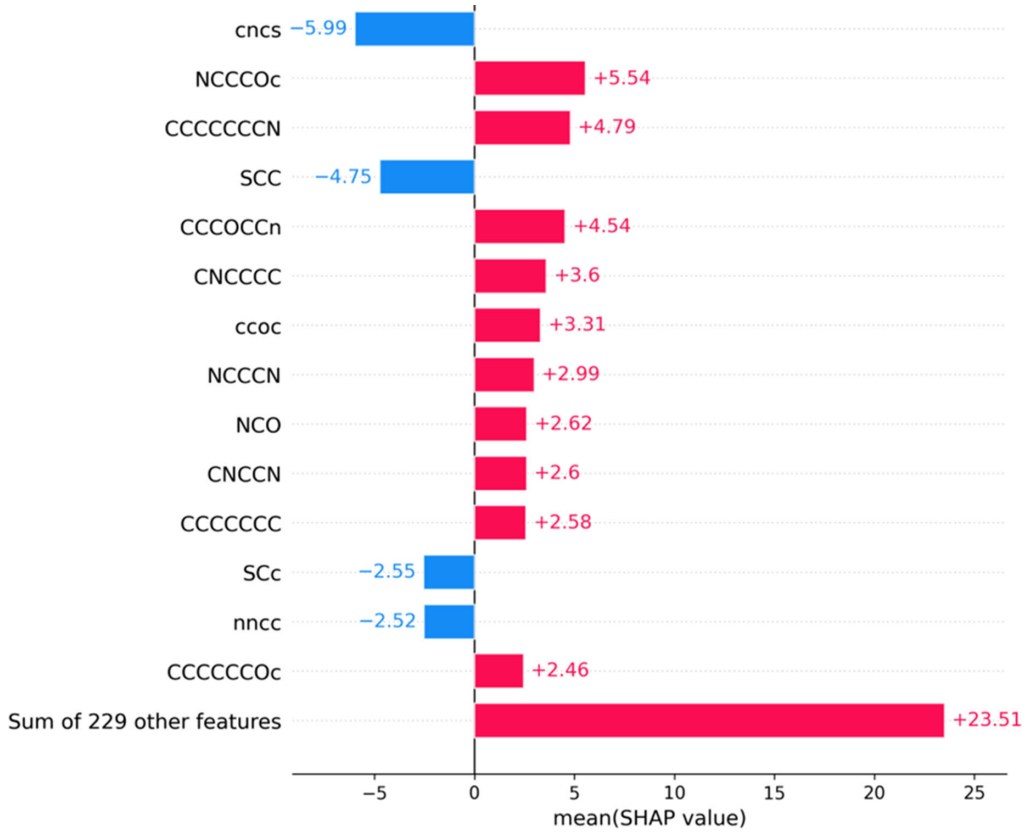

Figure 5. SHAP analysis

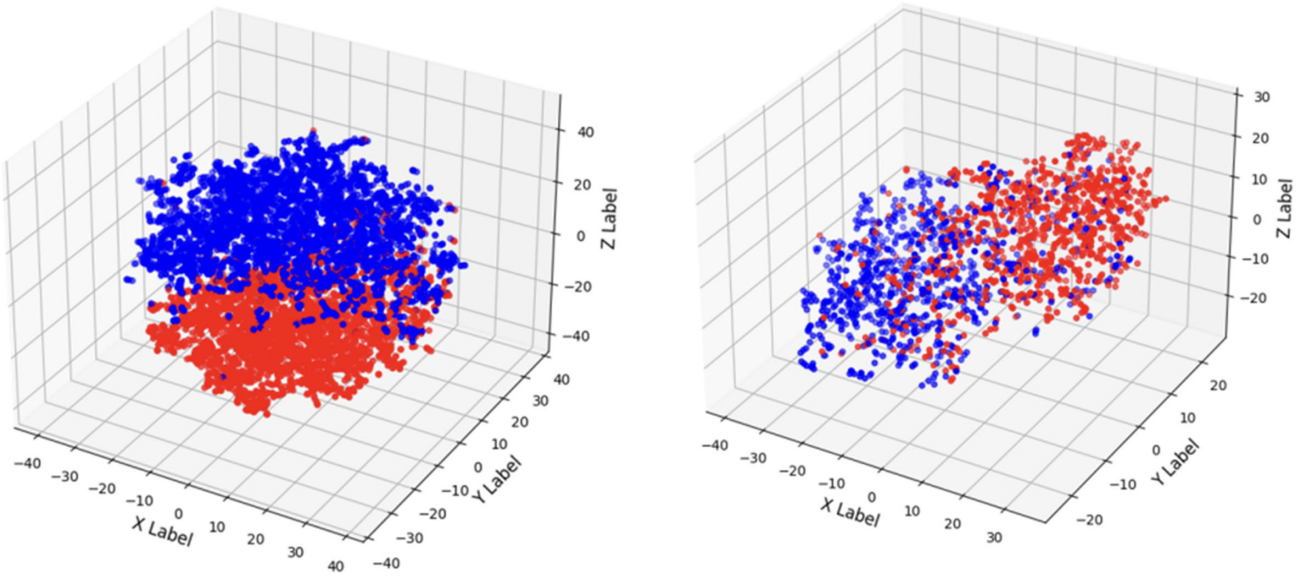

Figure 6. Representation visualization by *t*-SNE. The left is the training result. The right is the test result.

4. Conclusion

In this study, we introduce CLOP-hERG, a computational framework specifically designed for predicting drug-induced cardiotoxicity mediated by the hERG blockers. CLOP-hERG integrates cutting-edge deep learning techniques such as the pre-trained RoBERTa model and contrastive learning mechanisms to capture complex molecular features, significantly outperforming models based on traditional machine learning methods and feature engineering. Through a series of experiments, CLOP-hERG demonstrated superior predictive capabilities and robustness.

While the CLOP-hERG model has yielded promising results, it is not without its limitations. A significant challenge is the presence of huge parameters within the model, which leads to decreased efficiency

and slower processing times. To address this, we are considering the implementation of knowledge distillation as a means to simplify our model. Additionally, the current scope of the CLOP-hERG model is somewhat narrow, as it is specifically tailored for hERG blockers. In order to broaden its utility and enhance its predictive power, we plan to explore more advanced data augmentation techniques and deep learning architectures, adopting feature fusion strategies. Such an approach would enable the model to effectively combine and analyze different types of data, potentially improving its accuracy and robustness. Furthermore, recognizing the importance of multifunctionality in drug research, we aim to expand the datasets used for training our model. By integrating various datasets, we hope to extend the applicability of our model to a wider range of pharmaceutical compounds.

In the future, in addition to extending the application of CLOP-hERG to other types of drug safety assessments, we also plan to explore more advanced techniques, including multimodal dataset integration and transfer learning. Moreover, we anticipate that by further refining the interpretability of the model and enhancing the user interface, the tool will become more user-friendly and accessible to pharmaceutical researchers, thereby having a broader impact in the field of drug discovery. Overall, CLOP-hERG has showcased its potential value in enhancing the safety and efficacy of drug design, heralding the future direction of AI-assisted drug discovery in a certain.

Funding Support

This work was supported in part by the JSPS KAKENHI Grant Number JP23H03411 and JP22K12144, and the JST Grant Number JPMJPF2017. This work was also supported by JST SPRING, Grant Number JPMJSP2124.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The CLOP-hERG data that support the findings of this study are openly available at https://github.com/heshida01/CLOP-hERG. The Therapeutics Data Commons data that support the findings of this study are openly available at https://tdcommons.ai/, reference number [21].

References

[1] Vandenberg, J. I., Walker, B. D., & Campbell, T. J. (2001). HERG K^+ channels: Friend and foe. *Trends in Pharmacological Sciences*, *22*(5), 240–246. https://doi.org/10.1016/S0165-6147(00)01662-X

[2] Rampe, D., Roy, M. L., Dennis, A., & Brown, A. M. (1997). A mechanism for the proarrhythmic effects of cisapride (Propulsid): High affinity blockade of the human cardiac potassium channel HERG. *FEBS Letters*, *417*(1), 28–32. https://doi.org/10.1016/S0014-5793(97)01249-0

[3] Woosley, R. L. (1996). Cardiac actions of antihistamines. *Annual Review of Pharmacology and Toxicology*, *36*, 233–252. https://doi.org/10.1146/annurev.pa.36.040196.001313

[4] Zhou, Z., Vorperian, V. R., Gong, Q., Zhang, S., & January, C. T. (1999). Block of HERG potassium channels by the antihistamine astemizole and its metabolites desmethylastemizole and norastemizole. *Journal of Cardiovascular Electrophysiology*, *10*(6), 836–843. https://doi.org/10.1111/j.1540-8167.1999.tb00264.x

[5] Kumar, S., & Veer, K. (2023). Evaluation of current trends in biomedical applications using soft computing. *Current Bioinformatics*, *18*(9), 693–714. https://doi.org/10.2174/1574893618666230706112826

[6] Wang, Z., Quan, Z., Wang, Z. J., Hu, X., & Chen, Y. (2020). Text to image synthesis with bidirectional generative adversarial network. In *IEEE International Conference on Multimedia and Expo*, 1–6. https://doi.org/10.1109/ICME46284.2020.9102904

[7] Wang, L., Yang, X., Kuang, L., Zhang, Z., Zeng, B., & Chen, Z. (2023). Graph convolutional neural network with multi-layer attention mechanism for predicting potential microbe-disease associations. *Current Bioinformatics*, *18*(6), 497–508. https://doi.org/10.2174/1574893618666230316113621

[8] Chen, Y., Wang, Z., Wang, L., Wang, J., Li, P., Cao, D., . . . , & Sakurai, T. (2023). Deep generative model for drug design from protein target sequence. *Journal of Cheminformatics*, *15*(1), 38. https://doi.org/10.1186/s13321-023-00702-2

[9] Yu, L., Shi, Y., Zou, Q., Wang, S., Zheng, L., & Gao, L. (2020). Exploring drug treatment patterns based on the action of drug and multilayer network model. *International Journal of Molecular Sciences*, *21*(14), 5014. https://doi.org/10.3390/ijms21145014

[10] Rogers, D., & Hahn, M. (2010). Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, *50*(5), 742–754. https://doi.org/10.1021/ci100050t

[11] Ren, Z. H., You, Z. H., Zou, Q., Yu, C. Q., Ma, Y. F., Guan, Y. J., . . . , & Pan, J. (2023). DeepMPF: Deep learning framework for predicting drug-target interactions based on multi-modal representation with meta-path semantic analysis. *Journal of Translational Medicine*, *21*(1), 48. https://doi.org/10.1186/s12967-023-03876-3

[12] Shang, Y., Gao, L., Zou, Q., & Yu, L. (2021). Prediction of drug-target interactions based on multi-layer network representation learning. *Neurocomputing*, *434*, 80–89. https://doi.org/10.1016/j.neucom.2020.12.068

[13] Ogura, K., Sato, T., Yuki, H., & Honma, T. (2019). Support Vector Machine model for hERG inhibitory activities based on the integrated hERG database using descriptor selection by NSGA-II. *Scientific Reports*, *9*(1), 12220. https://doi.org/10.1038/s41598-019-47536-3

[14] Wang, T., Sun, J., & Zhao, Q. (2023). Investigating cardiotoxicity related with hERG channel blockers using molecular fingerprints and graph attention mechanism. *Computers in Biology and Medicine*, *153*, 106464. https://doi.org/10.1016/j.compbiomed.2022.106464

[15] Karim, A., Lee, M., Balle, T., & Sattar, A. (2021). CardioTox net: A robust predictor for hERG channel blockade based on deep learning meta-feature ensembles. *Journal of Cheminformatics*, *13*(1), 60. https://doi.org/10.1186/s13321-021-00541-z

[16] Liu, C., Wan, Z., Cheng, S., Zhang, M., & Arcucci, R. (2023). ETP: Learning transferable ECG representations via ECG-text pre-training. *arXiv Preprint:2309.07145*. https://doi.org/10.48550/arXiv.2309.07145

[17] Liu, C., Cheng, S., Chen, C., Qiao, M., Zhang, W., Shah, A., . . . , & Arcucci, R. (2023). M-FLAG: Medical vision-language pre-training with frozen language models and latent

space geometry optimization. In *Proceedings of 26th International Conference on Medical Image Computing and Computer-Assisted Intervention*, 637–647. https://doi.org/10.1007/978-3-031-43907-0_61

[18] Wan, Z., Liu, C., Zhang, M., Fu, J., Wang, B., Cheng, S., . . . , & Arcucci, R. (2023). Med-UniC: Unifying cross-lingual medical vision-language pre-training by diminishing bias. *arXiv Preprint:2305.19894*. https://doi.org/10.48550/arXiv.2305.19894

[19] Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J., & Bryant, S. H. (2009). PubChem: A public information system for analyzing bioactivities of small molecules. *Nucleic Acids Research*, *37*, W623–W633. https://doi.org/10.1093/nar/gkp456

[20] Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. In *Proceedings of the 37th International Conference on Machine Learning, 119*, 1597–1607.

[21] Huang, K., Fu, T., Gao, W., Zhao, Y., Roohani, Y., Leskovec, J., . . . , & Zitnik, M. (2022). Artificial intelligence foundation for therapeutic science. *Nature Chemical Biology*, *18*(10), 1033–1036. https://doi.org/10.1038/s41589-022-01131-2

[22] Kim, H., Park, M., Lee, I., & Nam, H. (2022). BayeshERG: A robust, reliable and interpretable deep learning model for predicting hERG channel blockers. *Briefings in Bioinformatics*, *23*(4), bbac211. https://doi.org/10.1093/bib/bbac211

[23] Landrum, G. (2010). *RDKit: Open-source cheminformatics software.* Retrieved from: https://www.rdkit.org

[24] Wright, R. E. (1995). Logistic regression. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 217–244). American Psychological Association.

[25] Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and Their Applications*, *13*(4), 18–28. https://doi.org/10.1109/5254.708428

[26] Kim, Y. (2014). Convolutional neural networks for sentence classification. *arXiv Preprint:1408.5882*. https://doi.org/10.48550/arXiv.1408.5882

[27] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[28] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv Preprint:1810.04805*. https://doi.org/10.48550/arXiv.1810.04805

[29] Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv Preprint:1609.02907*. https://doi.org/10.48550/arXiv.1609.02907

[30] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2017). Graph attention networks. *arXiv Preprint:1710.10903*. https://doi.org/10.48550/arXiv.1710.10903

[31] Kim, H., & Nam, H. (2020). hERG-Att: Self-attention-based deep neural network for predicting hERG blockers. *Computational Biology and Chemistry*, *87*, 107286. https://doi.org/10.1016/j.compbiolchem.2020.107286

[32] Goh, G. B., Hodas, N. O., Siegel, C., & Vishnu, A. (2017). SMILES2Vec: An interpretable general-purpose deep neural network for predicting chemical properties. *arXiv Preprint:1712.02034*. https://doi.org/10.48550/arXiv.1712.02034

[33] Shao, J. (2003). *Mathematical statistics*. USA: Springer & Business Media.

[34] Tallarida, R. J., & Murray, R. B. (1987). *Manual of pharmacologic calculations: With computer programs*. USA: Springer.

[35] Ramsundar, B., Eastman, P., Walters, P., & Pande, V. (2019). *Deep learning for the life sciences: Applying deep learning to genomics, microscopy, drug discovery, and more*. USA: O'Reilly Media.

[36] Lundberg, S. M., Nair, B., Vavilala, M. S., Horibe, M., Eisses, M. J., Adams, T., . . . , & Lee, S. I. (2018). Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nature Biomedical Engineering*, *2*(10), 749–760. https://doi.org/10.1038/s41551-018-0304-0

[37] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Proceedings of 31st Conference on Neural Information Processing Systems*, 1–10.

[38] Hinton, G. E., & Roweis, S. (2002). Stochastic neighbor embedding. In *Proceedings of the 2002 Neural Information Processing Systems Conference*, 1–8.

How to Cite: He, S., Ye, X., & Sakurai, T. (2024). CLOP-hERG: The Contrastive Learning Optimized Pre-trained Model for Representation Learning in Predicting Drug-Induced hERG Channel Blockers. *Medinformatics*, *1*(3), 103–111. https://doi.org/10.47852/bonviewMEDIN42022049