NADR-Net: A Deep Learning Framework for Predicting Neurological Adverse Drug Reactions Using 17 Molecular Descriptors

Anushka Chaurasia[1,*,†], Deepak Kumar[1,†] and Yogita[2,†]

[1]Department of Computer Science and Engineering, National Institute of Technology, Meghalaya, India
[2]Department of Computer and Engineering, National Institute of Technology Kurukshetra, India

Abstract

Neurological adverse drug reactions (NADRs) pose a significant clinical challenge as they can have a profound impact on patient health and treatment outcomes. While diverse drug descriptors have been employed for neurological ADR prediction, the potential of using 17 molecular descriptors for this purpose has not been explored. To address this, a multilabel NADR-Net and MLSMOTE-based framework have been proposed for neurological adverse drug reaction prediction. The data for 17 MD and ADRs were extracted from PubChem and ADRECs databases and then mapped based on drug ID. The resulting dataset contained information on 2160 drugs, including their molecular properties, and 1030 ADRs. The methodology was then applied to this dataset, and it showed promising results in terms of hamming loss, precision, true positive rate, f1 score, and ROC-AUC. This study highlights the potential of using molecular descriptors for predicting neurological ADRs, which could improve patient outcomes and drug safety.

Keywords

Neurological Adverse Drug Reaction, Deep Neural Network, 17 Molecular Descriptors (17MD), MLSMOTE

1. Introduction

An Adverse Drug Reaction (ADR) is characterized as a Superfluous or deleterious reaction experienced after the administration of a medication [1]. These reactions may vary in severity and have the potential to impact any organ or body system. The timely detection of ADRs is recognized as essential for averting further complications and ensuring patient safety. In recent times, there has been a noted increase in the prevalence of neurological adverse drug reactions (NADRs) [2]. A study conducted in Central India [3] focused on the analysis and presentation of the occurrence and severity of spontaneous ADR reports. It was observed in this study that the majority of ADRs, particularly neurological ADRs, manifested within the initial five days following medication commencement [4]. Additionally, another research identified over 47,000 ADR reports associated with metoclopramide use. Similar to the previous findings, it was noted

Symposium on Computing and Intelligent Systems (SCI), May 10, 2024, New Delhi, India

*Corresponding author.
†These authors contributed equally.
anushkachaurasia1601@gmail.com (A. Chaurasia); deepak.kumar@nitm.ac.in (D. Kumar); yogitathakran@nitm.ac.in (Yogita)
0000-0002-2176-1840 (A. Chaurasia)

that most ADRs, including neurological ones, typically occurred within the first few days of starting metoclopramide therapy [5] [6]. From the literature perspective, considerable research has been conducted in the field of ADR prediction. However, when it comes to the prediction of NADRs, only a limited number of studies have been undertaken. These studies [7] have primarily focused on employing biological, structural fingerprint, and phenotypic drug properties for the prediction of 22 neurological ADRs, framing it as a binary classification problem. They have notably utilized machine learning techniques, which necessitated the selection of relevant features. The challenge of NADR detection, characterized as a multilabel issue where a single drug can induce multiple ADRs concurrently, accentuates the importance of this research. Therefore, the main aim of this study is to predict NADRs using 17 molecular Descriptors(MD). These 17 MD are considered essential for the prediction of ADRs as they are intrinsically linked to a drug's physicochemical characteristics, which directly impact its pharmacokinetics (the movement of drugs within the body) and pharmacodynamics (the biological effects of drugs on the body).

To achieve this aim, a deep neural network-based framework, NADR-Net, is designed to process and analyze the 17 MD characteristics of drugs to identify potential neurological adverse effects. To handle class imbalance, MLSMOTE is applied. The efficacy of this framework is validated using five performance metrics: precision, true positive rate, f1-score, AUC-ROC score, and hamming loss. Furthermore, the performance of the proposed architecture is compared with two other DNN architectures, namely, the standard DNN and the DNN with a 1-skip connection. Subsequently, the framework's performance, with and without the use of MLSMOTE, is compared to demonstrate the influence of the MLSMOTE technique on the model's proficiency in accurately predicting NADRs.

The remainder of this paper is structured as follows: Section 2 presents related work, Section 3 outlines the dataset and proposed methodology for predicting NADR-Net, Section 4 details the experimental setup and results, and Section 6 offers concluding remarks.

2. Related Work

In the last few years, several proposals have been made to use machine learning and deep learning to address the problem of ADR prediction. One such study [7] employed a binary classification approach and evaluated 176 machine-learning models for 22 NADRs. The models were trained using drug properties, including biological, chemical, and phenotypic aspects. SMOTE was applied to mitigate class imbalance. Relief-based feature selection techniques were used to identify relevant drug properties. Lee et al. [8] introduced a three-interval method integrating chemical and biological drug properties, outperforming k-nearest neighbor, naïve Bayes, and random forest classifiers. Another study [9] suggested a hybrid clustering-based approach to analyze the quantitative relationships between adverse drug reactions and patient attributes. Further, Jamal et al. [10] focused on predicting adverse drug reactions for cardiovascular drugs through biological and chemical information, therapeutic indications, and their combined datasets, using techniques specifically RF, SVM, and Sequential Minimization Optimization. This approach addressed the issue of class imbalance through SMOTE. Another study describes an approach to generating DL-based, systematic ADR prediction models [11] [12]. In a related

study, Dey et al. [13] use ML models, including a DL framework, to simultaneously predict ADRs and identify the underlying mechanisms. The models were trained on chemical-protein binding and gene expression datasets to improve prediction performance. Zheng et al. [14] examined and pinpointed the adverse effects associated with medications through the use of Highly Credible Negative Samples (HCNS), which were derived from various sources, including pathways, target proteins, chemical substructures, substituents, and the connection between genes and diseases. Their research dataset included 1,048 drugs and 1,276 different side effects. Das et al. [15] explored a multi-label ML strategy using drug functions and the MLSMOTE technique for handling class imbalances. In a recent review article, Lee et al. [16] provides an overview of the detection and classification of side effects using deep learning approaches. They show that deep learning approaches can help reduce or prevent the occurrence of ADRs by detecting and predicting them during post-marketing surveillance. Martenot et al. [17] proposed a DL-based pipeline for ADR monitoring in the biomedical literature that was introduced to detect serious ADRs in relevant documents at the sentence level. It relies on a modular architecture of open-source fine-tuned models and drug entities. Wang et al. [18] developed a deep learning model to examine drug side effects using descriptors from various sources, including biomedical literature from MEDLINE, drug-like and biological properties from PubChem and DrugBank, respectively. These elements were combined into a dataset for analysis using an MLP model with two hidden layers. The model outperformed others like Gaussian NB, Linear SVM, and PMF, achieving a top AUC of 84.40

The investigation of NADRs remains primarily focused on chemical and biological properties in existing literature, with only one study [7]conducted in this domain. However, a notable lack of research using a set of 17 specific molecular properties to explore these ADRs exists. Prior research has employed traditional machine learning models and the SMOTE for data balancing, limiting their scope to 22 non-neurological ADRs (NADRs). To date, the potential of deep learning models in studying neurological ADRs remains largely unexplored, highlighting a significant gap in current research.

3. Dataset and Methodological Framework

This section presents the acquisition of 17 MD datasets to predict neurological adverse drug reactions. Furthermore, the problem statement for predicting NADR and the proposed framework for solving the stated problem has been demonstrated.

3.1. Problem Statement

Let $D = \{d_1, d_2, \ldots, d_D\}$ be a set of drugs, where D is the total number of drugs. Each drug d_i is characterized by a feature vector $\mathbf{x}_i \in \mathbb{R}^{17}$, representing the 17 molecular properties of the drug. The objective is to map each drug d_i to a set of neurological ADRs. Define the set of all possible ADRs as $A = \{a_1, a_2, \ldots, a_{1030}\}$, where each a_j represents a unique neurological ADR. The task is to learn a function $f : \mathbb{R}^{17} \rightarrow \{0,1\}^{1030}$ such that for each drug d_i, the function predicts a binary vector $\mathbf{y}_i = f(\mathbf{x}_i)$, where $\mathbf{y}_i \in \{0,1\}^{1030}$. In this vector, $y_{ij} = 1$ indicates the presence of the ADR a_j for the drug d_i, and $y_{ij} = 0$ indicates its absence.

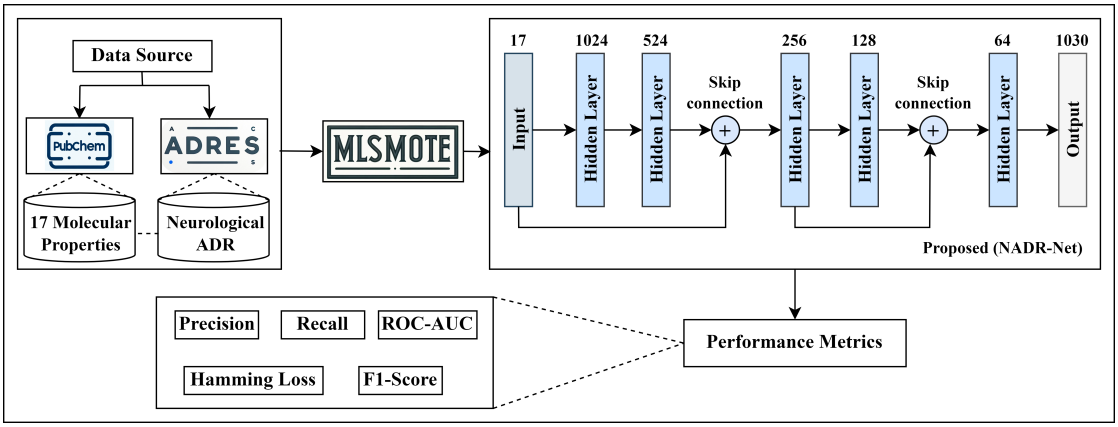


Figure 1: Block Diagram of the Proposed Framework for Neurological ADR Prediction.

3.2. Dataset Acquisition

In this section, the preparation of datasets for validating the proposed methodology is carried out by integrating 17 molecular drug descriptors and neurological adverse drug reactions (NADRs) data, as illustrated in Figure. 1. The properties of 17 molecules from drugs are retrieved from the PubChem [19] database using the PubChemPy Python package, resulting in a dataset of 2310 drugs, each characterized by 17 distinct features. These features include exact mass, hydrogen bond donor count, molecular weight, covalently-bonded unit count, rotatable bond count, undefined bond stereocenter count, complexity, monoisotopic mass, defined bond stereocenter count, topological polar surface area, isotope atom count, hydrogen bond acceptor count, formal charge, heavy atom count, defined atom stereocenter count, and XlogP3. The data on neurological adverse drug reactions, comprising 2160 drug samples and information on the occurrence and non-occurrence of 1030 ADRs, is extracted from the ADRECs [20] database. The dataset is created by mapping the 17 molecular drug descriptors with ADR data on drug ID.

3.3. Multilabel Synthetic Minority Over-sampling Technique (MLSMOTE)

In this study, we applied the MLSMOTE [21] to mitigate the challenge of underrepresented data in NADRs. MLSMOTE enhances the dataset by creating synthetic samples, specifically chosen for its ability to consider the multi-label characteristics of our data. The process begins with identifying minority labels, which are those with fewer average data samples compared to others. Subsequently, all data samples associated with these minority labels are identified. We then randomly select a sample from this minority group and identify its k-nearest neighbors to form a reference neighborhood. The generation of synthetic data samples involves an interpolation method for attributes and a majority voting approach for labels, utilizing the selected minority sample and its neighbors. In this context, the number of nearest neighbors (k) was set to 3.

3.4. Proposed Framework

Adverse neurological drug reactions are predicted using a deep neural network-based architecture (NADR-Net), as detailed in Table. 1. This architecture utilizes a sequential and concatenated approach to process 17 molecular properties. The model begins with an input layer, indicating the dataset's feature count, followed by dense layers. Each layer employs a ReLU (Rectified Linear Unit) activation function, introducing non-linearity and enhancing the model's ability to learn complex patterns.

Table 1
NADR-Net Architecture with Different Parameters

No.	Layer	Output Shape	Concatenate	Parameter
1	Input Layer	None,17	-	0
2	Dense	None,1024	Input layer	18,432
3	Dense_1	None,, 528	Dense	541,200
4	Concatenate	None,548	Input , Dense_1	0
5	Dense_2	None,256	Concatenate	139,776
6	Dense_3	None,128	Dense_2	32896
7	Concatenate_1	None,384	Dense_2, Dense_3	0
8	Dense_4	None,64	Concatenate	24640
9	Dense_5	None,1030	Dense_4	66950
Total Parameters				823,894
Trainable Parameters				823,894
Non-trainable Parameters				0

The input data is transformed into a compressed representation in the described neural network model through a series of dense layers. Initially, the input is compressed into a 1024-dimensional space by the first layer, which is then reduced to 528 dimensions by the second layer. Following this, the output of the second layer is concatenated with the original input, resulting in a 545-dimensional vector. This vector undergoes further dimensionality reduction through two additional dense layers, ultimately bringing it down to 128 dimensions. Another concatenation operation combines the output of the third layer with that of the second layer to leverage low-level and high-level features from the previous layers. Finally, the resulting vector is transformed by one last dense layer into a 64-dimensional output. For the final output, a 1030-dimensional vector is produced using a sigmoid activation function, indicating the model's capacity for multi-label classification. By combining earlier and deeper features in the network, this architecture enhances the learning process through feature reusability and representation learning.

4. Experimental Setup and Results

This section presents the evaluation measures, experimental setup, and findings from analyzing 17 molecular properties in predicting neurological adverse drug reactions using the proposed framework, NADR-Net.

4.1. Evaluation Metrics

The performance of the proposed framework was evaluated using the following measures. Consider $D_s = \{(X_{\text{test}_j}, TL_j)|j = 1, 2, \ldots, n\}$ as the multi-label dataset, where AL_j signifies the actual label set for the test instance X_{test_j}, and PL_j represents the labels predicted by the classifier.

4.1.1. Hamming Loss:

Is defined as the frequency at which the model inaccurately predicts a label pair for a given sample PL_j. In this context, Δ represents the symmetric difference between two sets, which is used to calculate the mismatch between the predicted and actual labels [22].

$$\text{Hamming Loss} = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{|AL|}|PL_j \Delta TL_j| \tag{1}$$

4.1.2. Precision:

Is the ratio of actual positive predictions to the total number of positive predictions made by the model [22].

$$\text{Precision} = \frac{1}{n}\sum_{j=1}^{n}\frac{|PL_j \cap AL_j|}{|PL_j|} \tag{2}$$

4.1.3. True Positive Rate (TPR):

Is defined as the ratio of the number of positive samples correctly identified by the model (True Positives) to the total number of actual positive samples in the data (the sum of True Positives and False Negatives) [22].

$$\text{Precision} = \frac{1}{n}\sum_{j=1}^{n}\frac{|PL_j \cap AL_j|}{|AL_j|} \tag{3}$$

4.1.4. F1-Score:

Is the harmonic average of TPR and Precision [22].

$$\text{F1 Score} = \frac{1}{n}\sum_{j=1}^{n}\frac{2 \times |\text{PL}_i \cap \text{AL}_i|}{|\text{PL}_i| + |\text{AL}_i|} \tag{4}$$

4.1.5. ROC-AUC score:

It indicates a model's ability to discriminate across classes, with the ROC curve plotting true versus false positive rates and the AUC quantifying the overall classification accuracy. Higher AUC indicates better model performance.

4.2. Experimental Setup

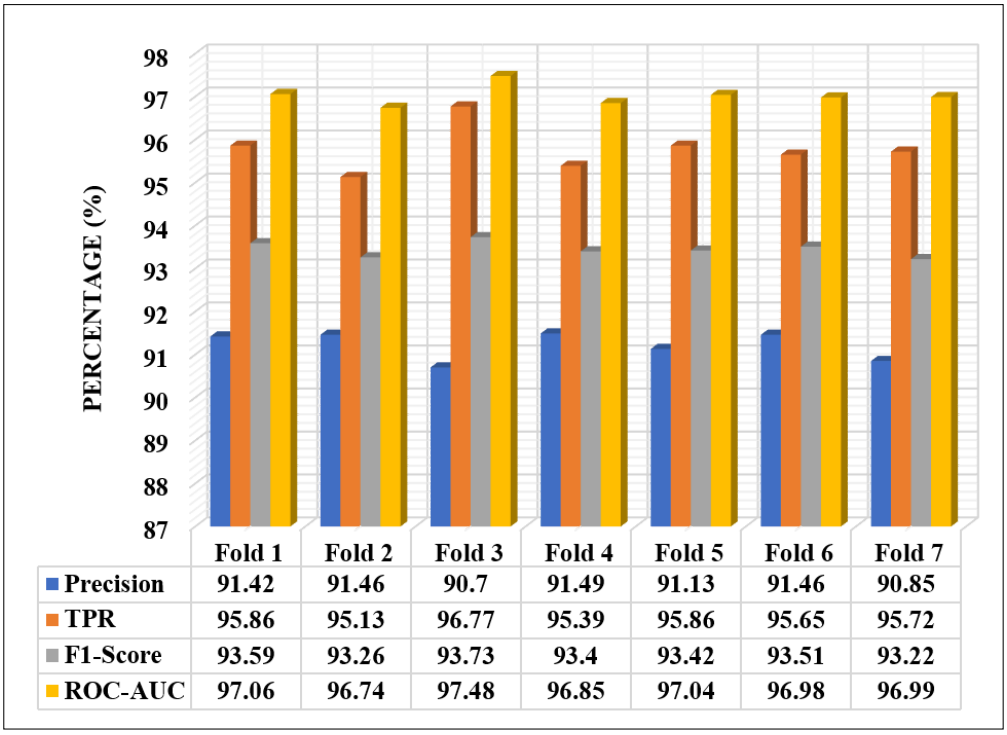
The implementation of the proposed framework was carried out using the Scikit-learn package in Python 3.7.4. An Intel (R) i5-5300U CPU and 512 GB of RAM were utilized by the deep learning framework. In the DNN architecture, the ReLU activation function and six hidden layers (1024, 528, 256, 128, 64, 32) were employed. For the DNN with one skip connection, two hidden layers (1024, 528) were used in conjunction with the ReLU activation function, followed by a concatenation layer that combines the input layer with the second hidden layer (528). Subsequently, four hidden layers (256, 128, 64, 32) were utilized. ReLU and Sigmoid activation functions were adopted to address non-linearities and probabilistic outputs, crucial for multi-label classification. A learning rate of 0.01 was set to ensure efficient convergence. Binary cross-entropy was selected as the loss function due to its effectiveness in independent label predictions. The Adam optimizer was used for its capability to manage sparse gradients. The training was conducted over 150 epochs with a batch size of 32, aiming to balance learning efficiency and prevent overfitting, optimizing the model's performance for the multi-label dataset. These hyperparameters were kept consistent across all three architectures. Additionally, the cross-validation technique was used for validation, where 80% of the data was allocated for training and 20% for testing.

4.3. Experimental Results and Discussion

This section analyses the effectiveness of the proposed framework, comparing it with DNN and DNN with 1-skip connection architectures, and evaluates the impact of using the MLSMOTE technique. The model is evaluated on test data and calculates various performance measures. Table 2 presents a comparative performance analysis of deep neural network, DNN with a 1-skip connection, and NADR-NET. Each architecture's performance is evaluated based on various metrics: precision, actual positive rate (TPR), f1-score, ROC-AUC, and hamming loss. Precision showed an incremental increase from 89.06% to 91.24%, TPR improved from 93.12% to 95.77%, and the f1-Score rose from 92.01 to 93.45. ROC-AUC and hamming loss saw similar trends, with ROC-AUC escalating from 95.65% to 97.02% and loss decreasing from 2.32 to 2.12. These results suggest that NADR-NET architecture significantly enhances model efficacy with increments ranging from approximately 2% to 5% across five metrics.

Table 2
Comparative Performance Analysis of Different Neural Network Architectures

Metrics	DNN	DNN+1-skip connection	NADR-NET
Precision(%)	89.06 ± 0.0067	90.23 ± 0.0032	91.24 ± 0.0028
TPR(%)	93.12 ± 0.0043	94.15 ± 0.0154	95.77 ± 0.0051
F1-Score	92.01 ± 0.0040	92.54 ± 0.0042	93.45 ± 0.0017
ROC-AUC	95.65 ± 0.0022	96.11 ± 0.0032	97.02 ± 0.0023
Hamming Loss	2.32 ± 0.0022	2.25 ± 0.0016	2.12 ± 0.0004



	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7
Precision	91.42	91.46	90.7	91.49	91.13	91.46	90.85
TPR	95.86	95.13	96.77	95.39	95.86	95.65	95.72
F1-Score	93.59	93.26	93.73	93.4	93.42	93.51	93.22
ROC-AUC	97.06	96.74	97.48	96.85	97.04	96.98	96.99

Figure 2: Performance Metrics of Proposed NADR-Net Framework on Each Fold.

As delineated in Figure. 2, precision, true positive rate (TPR), f1-score, and ROC-AUC are presented for each fold, ranging from 1 to 7. Precision percentages fluctuate slightly, maintaining above 90% across all folds. TPR predominantly remains in the mid-95% range, peaking at 96.77% in fold 3. The F1-Score, indicative of test accuracy, consistently stays above 93%, reflecting robust model performance. ROC-AUC values, representing the model's class distinction ability, are uniformly high, with all folds scoring above 96%. It can be observed from the above graph that the predictive model exhibits a high degree of accuracy and reliability, as evidenced by the performance metrics across all folds in the cross-validation process.

The incorporation of MLSMOTE was evaluated both with and without the NADR-Net's performance in the research study. The results, as demonstrated in Figure. 3, indicated a substantial enhancement in the model's performance. More than a 42% increase was shown in precision, while the rate of correctly identified positive instances experienced a growth of over 55%. A nearly 49% improvement was exhibited in the aggregate measure of test accuracy, and the model's ability to differentiate between classes improved by approximately 28%. Furthermore, the occurrence of incorrect labels decreased by over 50%, underscoring the significant impact of MLSMOTE on the overall effectiveness of the NADR-Net.

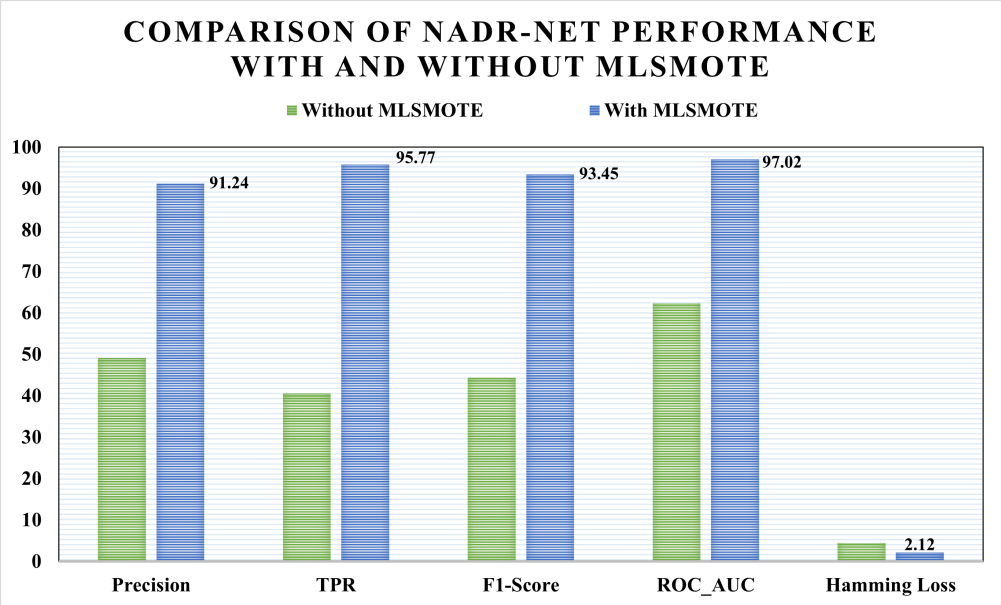


Figure 3: Comparison of NADR-Net performance with and without Augmentation

5. Conclusion

The proposed methodology effectively showcased the capability of the NADR-Net, a deep neural network, to predict adverse neurological drug reactions utilizing 17 molecular properties. A significant aspect of the study was addressed by applying the Multilabel Synthetic Minority Over-sampling Technique (MLSMOTE) to overcome the challenge of class imbalance. The efficacy of NADR-Net was evaluated rigorously using five performance metrics: a high precision of 91.24 ± 0.0028, TPR of 95.77 ± 0.0051, F1 score of 93.45 ± 0.0017, ROC-AUC of 97.02 ± 0.0023 and maintaining a hamming loss at a minimal 2.12 ± 0.0004 were notably achieved. A comparative analysis revealed that the model's performance was notably enhanced by including the MLSMOTE-enhanced dataset. In the future, an expanded set of drug properties could be utilized to enhance prediction capabilities and integrate mechanisms to demystify the black box model's decision-making process.

References

[1] K. Patton, D. Borshoff, Adverse drug reactions, Anaesthesia 73 (2018) 76–84.
[2] K. K. Jain, K. K. Jain, Neurologic symptoms as adverse drug reactions, Drug-induced Neurological Disorders (2021) 11–25.
[3] Adverse drug reactions - behind the headlines, https://www.bhf.org.uk/informationsupport/heart-matters-magazine/news/behind-the-headlines/adverse-drug-reactions, 2023. Accessed: 2024-01-19.
[4] M. Sharma, R. Baghel, S. Thakur, S. Adwal, Surveillance of adverse drug reactions at an

adverse drug reaction monitoring centre in central india: A 7-year surveillance study, BMJ open 11 (2021).
[5] K. Svendsen, M. Wood, E. Olsson, H. Nordeng, Reported time to onset of neurological adverse drug reactions among different age and gender groups using metoclopramide: an analysis of the global database vigibase®., European Journal of Clinical Pharmacology 74 (2018) 627–636.
[6] P. A. Thuermann, R. Windecker, J. Steffen, M. Schaefer, U. Tenter, E. Reese, H. Menger, K. Schmitt, Detection of adverse drug reactions in a neurological department: comparison between intensified surveillance and a computer-assisted approach, Drug safety 25 (2002) 713–724.
[7] S. Jamal, S. Goyal, A. Shanker, A. Grover, Predicting neurological adverse drug reactions based on biological, chemical and phenotypic properties of drugs using machine learning models, Scientific reports 7 (2017) 872.
[8] W.-P. Lee, J.-Y. Huang, H.-H. Chang, K.-T. Lee, C.-T. Lai, Predicting drug side effects using data analytics and the integration of multiple data sources, IEEE Access 5 (2017) 20449–20462.
[9] Yogita, J. W. Sangma, S. N. Anal, V. Pal, Clustering-based hybrid approach for identifying quantitative multidimensional associations between patient attributes, drugs and adverse drug reactions, Interdisciplinary Sciences: Computational Life Sciences 12 (2020) 237–251.
[10] S. Jamal, W. Ali, P. Nagpal, S. Grover, A. Grover, Computational models for the prediction of adverse cardiovascular drug reactions, Journal of translational medicine 17 (2019) 1–13.
[11] A. Mohsen, L. P. Tripathi, K. Mizuguchi, Deep learning prediction of adverse drug reactions in drug discovery using open tg–gates and faers databases, Frontiers in Drug Discovery 1 (2021) 768792.
[12] P. Das, Yogita, V. Pal, Integrative analysis of chemical properties and functions of drugs for adverse drug reaction prediction based on multi-label deep neural network, Journal of Integrative Bioinformatics 19 (2022) 20220007.
[13] S. Dey, H. Luo, A. Fokoue, J. Hu, P. Zhang, Predicting adverse drug reactions through interpretable deep learning framework, BMC bioinformatics 19 (2018) 1–13.
[14] Y. Zheng, H. Peng, X. Zhang, Z. Zhao, J. Yin, J. Li, Predicting adverse drug reactions of combined medication from heterogeneous pharmacologic databases, BMC bioinformatics 19 (2018) 49–59.
[15] P. Das, Y. Thakran, S. N. Anal, V. Pal, A. Yadav, Brmcf: Binary relevance and mlsmote based computational framework to predict drug functions from chemical and biological properties of drugs, IEEE/ACM Transactions on Computational Biology and Bioinformatics (2022).
[16] C. Y. Lee, Y.-P. P. Chen, Prediction of drug adverse events using deep learning in pharmaceutical discovery, Briefings in Bioinformatics 22 (2021) 1884–1901.
[17] V. Martenot, V. Masdeu, J. Cupe, F. Gehin, M. Blanchon, J. Dauriat, A. Horst, M. Renaudin, P. Girard, J.-D. Zucker, Lisa: an assisted literature search pipeline for detecting serious adverse drug events with deep learning, BMC medical informatics and decision making 22 (2022) 1–16.
[18] C.-S. Wang, P.-J. Lin, C.-L. Cheng, S.-H. Tai, Y.-H. Kao Yang, J.-H. Chiang, Detecting potential adverse drug reactions using a deep neural network model, Journal of medical

Internet research 21 (2019) e11016.
[19] S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, et al., Pubchem substance and compound databases, Nucleic acids research 44 (2016) D1202–D1213. doi:10.1093/nar/gkv951.
[20] M.-C. Cai, Q. Xu, Y.-J. Pan, W. Pan, N. Ji, Y.-B. Li, H.-J. Jin, K. Liu, Z.-L. Ji, Adrecs: an ontology database for aiding standardization and hierarchical classification of adverse drug reaction terms, Nucleic acids research 43 (2015) D907–D913.
[21] F. Charte, A. J. Rivera, M. J. del Jesus, F. Herrera, Mlsmote: Approaching imbalanced multilabel learning through synthetic instance generation, Knowledge-Based Systems 89 (2015) 385–397.
[22] D. Krstinić, M. Braović, L. Šerić, D. Božić-Štulić, Multi-label classifier performance evaluation with confusion matrix, Computer Science & Information Technology 1 (2020).