International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 07 | Jul 2026 www.irjet.net p-ISSN: 2395-0072

A MULTI-MODAL DEEP LEARNING FRAMEWORK FOR DRUG SIDE EFFECT FORECASTING

[1]Palle Sharath, [2]G. Raju

[1]M.Tech Scholar, Malla Reddy University, Hyderabad
[2]Associate Professor, Malla Reddy University, Hyderbad

ABSTRACT-*Quick and precise anticipation of the adverse drug reactions (ADRs) is critical towards enhancing pharmaceutical safety, minimizing clinical trial failures and facilitating effective pharmaceutical development. The research paper provides DeepSide, a multi-modal deep learning model that can be used to forecast drug side effects by using chemical structure features with SMILES representations, gene expression signatures, and experiment-specific metadata. In contrast to classical machine learning models, which use few datasets or manually-created features, DeepSide uses all the LINCS L1000 gene expression data and the below-the-hook architectures, MLP, Residual MLP, Multi-Mode Neural Networks, Multi-Task Neural Networks, as well as a SMILES-based CNN, to automatically learn complex drug-response behavior. The framework is a high-dimensional, heterogeneous, multi-label, multi-label classification of high- dimensional biomedical inputs to determine potential ADRs with increased accuracy. The experimental findings indicate that DeepSide is highly competitive with state-of-the-art models, and the performance on macro-AUC and micro-AUC indicators suffices to note a significant improvement, and new biologically meaningful side-effect associations that have not been recorded in existing databases are identified by DeepSide. Combining different biomedical modalities and leveraging deep neural representations, DeepSide has a great potential as a smart computational method to early toxicity screening, drug repurposing, and improved pharmacovigilance.*

Keywords: ADR, Deepside, MLP, smiles, AUC, toxicity, DL, CNN.

I INTRODUCTION

The prediction of adverse drug reactions (ADRs) is one of the most challenging and important processes of the ongoing pharmaceutical research because occasionally the unexpected side effects can be lethal to the patients who have been introduced to a clinical trial, costly failure of the experiment may occur, and the loss of money may be significant. The traditional in-vitro and in-vivo toxicity tests approach is not comprehensive due to its high cost, time-consuming nature of development, and absence of biological coverage; hence may not be utilized to identify any unusual or complex drug reactions. As the chemical libraries and high-throughput screening technology have been rapidly expanding, the need to have effective computational models that can be exploited to predict ADRs before clinical exposure has also been increasing. Consequently, the concept of deep learning and artificial intelligence will find its applications to improve experimental testing and augment the degree of drug safety testing during the initial stages.

Recent advances in deep learning have enabled models to be useful and learn about high-dimensional biomedical data without learning features that are designed by hand. Convolutional/fully connected neural networks can now transform the properties of drugs (e.g. chemical structure, molecular fingerprints or the SMILES sequence) into rich latent representations. In addition, very informative data on cell responses to drugs can be provided by vastly expressed datasets of genes (even those generated through the LINCS L1000 platform), and context-dependent biological signature can be found by traditional models. Integrating such heterogeneous sources of data has given the chance to enhance the precision of ADR forecasting that relies on multi-modal learning systems.

However, the existing computational ADR prediction algorithms are restricted to a few problems, including the use of limited high quality data sets, and failure to analyze dirty/diverse biological conditions, and multi- label classification. The majority of traditional models just base on one experiment of the drug and the fact that the vast bulk of available expression signatures of genes are ignored and most crucially of all, the vital contextual differences (dose, exposure time, and cell-line specificity) are ignored. Similarly, rule or similarity-based machine learning is not applicable to the non-linearity of drug properties and side effects and hence cannot be used effectively and predict unobserved compounds.

The DeepSide framework offers a framework of multi-modal prediction using deep learning including chemical structures, SMIles, gene expression profiles, and experimental metadata features into one system in order to overcome these. DeepSide can be used to learn the intricate patterns in multi-dimensional data, and learns the biological complexity of cellular responses to drugs using architectures such as Multi-Layer Perceptron (MLP), Residual MLP, Multi- Mode Neural Networks (MMNN), Multi- Task Neural Networks (MTNN) and SMILES- based Convolutional Neural Networks (SMILESConv). The multi-source fusion is known to not only enhance the accuracy of

prediction, but also enables significant ADR associations to be found not possible in traditional datasets.

Lastly, DeepSide can significantly contribute to the sphere of computational drug safety and offers a scalable and data- driven early toxicity screening method, drug repurposing and pharmacovigilance. The framework can provide a highly influential platform that can be more reliable in the identification of common and uncommon side effects since the full list of the LINCS L1000 repository is utilized and the recent developments in deep learning algorithms to detect them. The models of ADR prediction

like DeepSide will become extremely important in the process of safer development of drugs, reducing the risks of clinical activities, and accelerating the process of finding effective therapeutic compounds since pharmaceutical research is moving further to precision medicine and personalized therapeutics.

II SURVEY OF RESEARCH

Sharan and Atias (2011) -An inference of network-based side-effects. In Atias and Sharan, it is suggested that the graph-based approach is used, the relation of drug-side-effects can be described in the form of coupled networks, drug-side-effect bipartite and drug-protein interaction networks. Their algorithm spreads data of annotation information throughout the network to accurately anticipate the probable ADRs of the under annotated drugs, and they utilize the truth that the drugs that trigger corresponding proteins are also likely to produce comparable adverse impacts. More to the point, however, this network diffusion and canonical- correlation-type algorithm is also improved with the fact that it uses topological nearness and does not rely on chemical closeness, permitting the inferences to be made indirectly by exploiting the neighbors of the multi- relational space. The article shows that it is robust in hundreds of drugs and the network formulations possess scalability and interpretability advantages. Its input would be conceptualization, ADR prediction as aproblem on heterogeneous biomedical graphs and this would be inspired further by deep and graph-based models which would not need to rely on engineered features and learn representations on multi-relational structures. Huang et al. (2011): Integration of systems pharmacology of drug safety.

Huang and co-authors suggest integrative systempharmacology framework, a combination of drug-target interaction, protein-protein interaction network, molecular annotation and curated side effects data. They combine network modules and functional annotations with particular ADRs by integrating drug targets into global interactomes and extracting extended network neighborhoods. Systems-level perspective provides indirect and pathway-level effects which are not accessible to single-target analyses, and can more effectively predict complex, organ-system toxicities. The method emphasizes that context is essential and not target identity to predict the safety of drugs in the context of cellular networks where the targets are encapsulated. The paper provides the reasoning of a combination of biological context (e.g., gene expression, pathways) to chemical characteristics to provide a more accurate model of the mechanisms that lead to the adverse outcomes are achieved in later deep-learning applications such as DeepSide. [3] Liu et al. (2012) - Massive ADR prediction on multi-type features.

Liu et al. elaborate on the fact that a machine- learning pipeline, working on a mass scale, can include chemical fingerprints, target proteins, pathway annotations, drug indications, as well as phenotype data to predict ADR, which is achieved through feature-length vectors. They tested their models (logistic regression, SVMs, random forests) on thousands of drugs and side effects and demonstrate that phenotypic properties and prior knowledge of ADR outperform those based on either chemistry or target lists. It is notable that they possess a firmly established practical framework that is in the past and signals on the instance of grave ADRs of withdrawn drugs. The paper introduces the convention of multi-source integration and it is easy to discern the drawbacks of mono-modality one-source strategies that encourages automated representation learning strategies, including deep multimodal networks, to be capable of accepting heterogeneous input, without necessarily undergoing extensive feature engineering.

Side effects in the article of Mizutani et al. (2012) are associated with drugs-protein interaction modules.

These are correlations of drug binding specifics of hundreds of proteins with large ADR libraries and Mizutani et al. apply sparse canonical correlation analysis to retrieve module of protein-side-effects. They do have some consistent sets of targets, which are correlated with particular adverse phenotypes which provide mechanistic theories of toxicity development according to their findings. The modular view points out that the joint effects of disturbed targets (rather than single-target effects) are typical of complex ADRs and that models of prediction ought to have the ability to cause higher-order patterns of interactions. The paper has provided a methodological way of converting structured data on biological interactions to informative features, a notion that is opposite to multimodal fusion in DeepSide in which an interaction-learned representation might be limited or expanded by an interaction-learned module to be more easily interpreted mechanistically.

Yamanishi et al. (2012) -Designing chemical and biological similarity space.

Yamanishi and others believe that side-effect descriptions can be described with the aid of a multi-space paradigm, with structural similarity of chemical compounds, biological similarity (target-based), and a kernel. Their framework builds structural and functional similarity of drugs by the complementary similarity kernel construction and learning mappings to ADR labels and is more predictive than single-space models. The ability to survive on incomplete information is also one of the strengths: the model can also transfer the knowledge on similarity relations in the cases when some modalities of data are not available. Multi-view learning is directly impacted on multi-modal neural settings such as DeepSide, where chemical fingerprints, expression data and metadata are co-learned in the current case, deep networks learn non-linear cross-space mappings, rather than relying on hand-crafted kernels.

Wang, Clark and Maayan (2016) ADR prediction LINCS L1000 Signatures. The first to use the LINCS L1000 gene expression compendium to predict ADRs are Wang et al., but the side effects are predicted at scale using the enrichment features of chemical fingerprints of expressions. Their portal and pipeline of SEP-L1000 indicate that drug- perturbed transcriptional responses are more accurate in predictions than the baselines based on chemistry only. In practice to minimize noise, however, they only perform one good quality experiment per compound throwing much of the breadth of LINCS away. The constraint is what encourages the DeepSide method of training deep models on the entire experimental set to learn within a dose-cell- time-point space to learn context dependent toxic responses that might otherwise be missed by the single experiment method.

Dey et al. (2018) -- Deep learning between ADRs and substructures.

According to Dey et al., a deep-learning architecture is reported and it can be interpreted to designate molecular substructures to predict toxicity and this gives the saliency-type explanations of how chemical fragments are linked to certain ADRs.

Their model learns internal properties that can be applied to substructural motifs and does not require predefined fragment libraries and, hence, allows chemists to reason about structural variations to take risk off. The majority of the conventional baselines are beneath the performances and contain actionable interpretable results. This contribution is orthogonal to DeepSide in that DeepSide can use chemical and expression data to make predictions but they can be better by including interpretability mechanisms like saliency or attention to enable them to understand what molecular patterns or expression patterns are driving the predictions to be more reliable and translational in medicinal chemistry and regulatory biology.

III WORKING METHODOLOGY

The methodology of the working process is initiated by the intensive data gathering, when several biomedical sources are gathered and integrated to form a solid reservoir of data on adverse drug reaction (ADR) prediction. The most important data parts are chemical structure fingerprints, SMILES of drug molecules, clinical side- effect labels of drugs in SIDER database, and high-dimensional gene expression pattern of LINCS L1000 project. These datasets are combined with a constant set of identifiers and experimental metadata, including dosage, type of cell lines, and exposure time, are also tied to make sure that biological context is maintained. Such a multi-source data pool is the basis of the construction of a credible deep learning-based ADR prediction system since it incorporates both the chemical and biological aspects of the behavior of drugs.

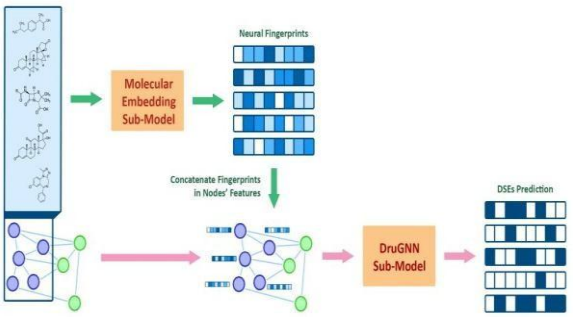

Fig.1. System architecture diagram

Preprocessing will be carried out after the inputs have been collected to transform the raw data into machine learning procedures. The gene expression profiles are made normalized to counter the difference in scale and remove experimental noise and metadata of categorical nature are one-hot coded to convert these into model ready numerical data. Multi-hot Vectors are suggested whereby the side-effect labels are coded as multi-hot label to achieve multi-label classification, in which a drug has a large number of potential ADRs. The issues of the missing values are addressed and the issues of the class imbalance is addressed with weighted loss functions to avoid the bias of the model to the side effects used most of the time. The strings of SMILES are padded and tokenized to correspond to the number of sequences in order to perform the convolutional processing easily. This preprocess pipeline guarantees the clean, regular and structured inputs and the deep neural structures.

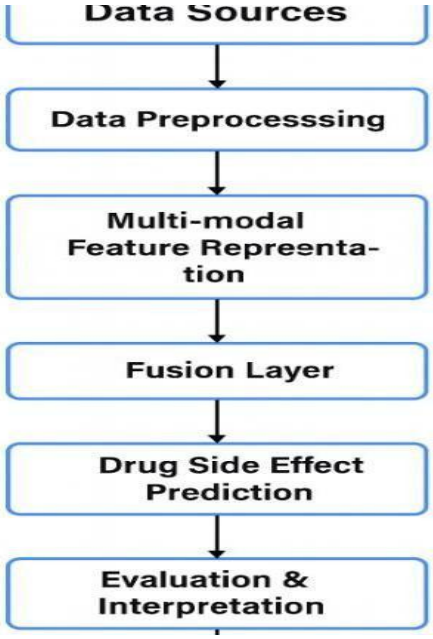

Fig.2. Flow of working

The second step goes by the multi- modal representation of features, chemical, biological and experimental features are organized in different and complementary streams of input. The advantage of this multi- modal approach is that it is capable of modeling drug behavior on different perspectives, which are structural, functional and contextual and consequently enhances the predictive richness of the system. The model is able to learn complex relations among these different modalities successfully by constructing citations of structure of representations that are tensors.

The approach is based on the deep learning framework development and training. They are estimated with a number of architectures applied to ADR prediction including Multi-Layer Perceptron (MLP), Residual MLP (ResMLP), Multi-Mode Neural Networks (MMNN), Multi-Task Neural Networks (MTNN) and SMILESConv based Convolutional Neural Networks. Each architecture is trained using best practices such as the use of batch normalization, dropout, ReLU activation and Adam optimizer. To address the imbalance between classes, cross- entropy loss is applied with weights; the imbalance between classes is corrected, and k- fold cross-validation is applied to ensure that the performance is measured in a robust manner. The SMILESConv architecture is superior among these models since it directly learns chemical substructures on tokenized SMILES sequences, and does not need handcrafted descriptors, and predicts better.

Finally, an optimized trained model is tested and applied into practice. Model compression techniques that could be used to reduce the inference time and computation overheads include pruning, quantization, and mixed-precision computation. The best model is deployed into deployable forms and integrated into an interface or REST API to approximate real-time ADR. The system is measured using Macro-AUC, micro-AUC, Hamming Loss, and mean Average Precision (mAP). The case-based analysis is done to confirm the prediction done by the model i.e. false positives are confirmed with the assistance of the independent biomedical literature. The final thing is to ensure that the DeepSide structure is not provided just with high predictive accuracy, but with meaningful and actionable insight on drug discovery, repurposing, and pharmacovigilance.

IV IMPLEMENTATION RESULTS EXPLANATION

DeepSide is activated by launching a massively large data integration chain that transforms a number of biomedical data sets into a framework of coherence that can be applied in deep learning. This is in the form of retrieval of LINCS L1000 gene expression profiles that provide high dimensional transcriptional responses to various doses, time exposures and cell lines. The clinical adverse drug reactions are also mapped to drug identifiers with the use of the SIDER database. The chemical structure information is acquired in the MACCS fingerprint form and the SMILES string form in order to have the molecular characterization exhaustively. The identifiers of non-present and inconsistent datasets are fixed by standardized drug reference lists and the appropriate correspondence of the chemical, genomic, and side-effect data is established throughout the integration period. The gene expression samples are filtered to remove any incomplete or unclean experiment and metadata, such as exposure time and cell line are numerically coded. The resultant joint collection of data contains thousands of labeled drug-experiment pairs and allows one to perform intensive multi-modal learning. This base is important because only clean, structured and high quality inputs can enabledeep neural networks to make accurate predictions. Combining chemical structure, biological response signature, and the experimental context, the performed system will permit the representation of drug behavior in the multidimensional nature and precondition the appropriate adverse drug reaction (ADR) modeling.

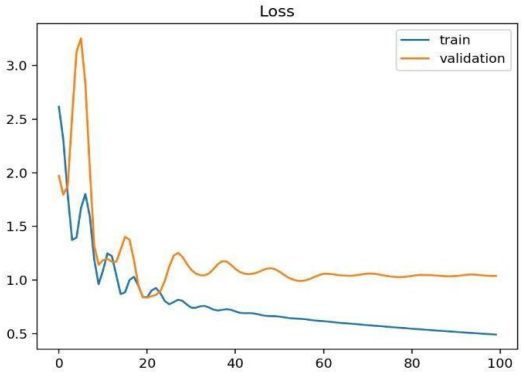

Fig.3. Traning dataset and validation

After the dataset has been prepared, there are multi-modal feature engineering and construction of tensors, which allow the deep learning models to be in position to accept heterogeneous inputs. Gene expression scales are converted to z-score or min-max scale to provide numerical stability and to avoid the domination of features that are of unequal magnitude. One-hot-encoded metadata values such as dose, time and cell line are converted into structured binary matrices with the set of categorical metadata values. A set of chemical structure fingerprints in the form of 166-bit MACCS vectors are transformed into fixed- length binary tensors, and SMILES strings into character tokenizations and character padded to the same length to allow convolutional operations. Such multi-modal illustrations makes the system recollect the global and local chemical data in addition to the biological setting. Combined features tensors have been organized in the form of different streams of the input that are entering different arms of the multi-modal neural structures during training. Close balancing will be exercised to balance the multi-label nature of ADRs wherein a drug may be linked to a lot of side effects. In order to reduce the label bias and the ADRs of rare type being identified, the weighted binary cross-entropy loss is used. This phase allows the learning models to be fed with noise- filtered, wide, and biologically pregnant inputs that have the capability of generating potent predictive tendencies throughout the training process.

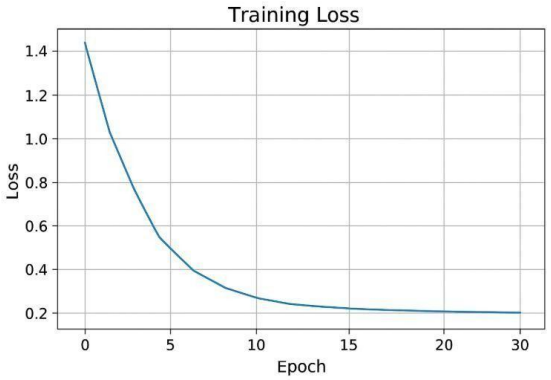

Fig.4. Training Loss in dataset

The main implementation procedure will be aimed at the training of various deep learning models that will help to define the degree of their effectiveness in drug side- effects prediction. Multi-Layer Perceptron (MLP) is built on the union of chemical and gene expression vectors with dropout layers and batch normalization in enhancing generalization. An even less specific variant of MLP is a Residual MLP (ResMLP), that has skip connections that can permit more elaborate network structures owing to the effect it has on vanishing gradients. The more complex design is Multi-Modal Neural Networks (MMNN) in which chemical, gene expression and metadata modalities are handled separately and then fused together in concatenation or summation fusion layers. Multi-Task Neural Network (MTNN) has the shared and task-specific output layers which are founded on the hierarchy of the ADR categories which allow the study to learn in a more biologically contextual manner. The SMILESConv model uses 1D convolutional filters of different kernel sizes to ensure that local motifs as well as long- range dependencies of molecular sequences have been discovered. Each of the models is trained on Adam optimizer, ReLU activation, mini- batches and weighted loss functions. The cross-validation is employed to make sure that there is generalization of the drugs that are not observed. Architectures are compared using performance measures like macro-AUC, micro- AUC and Hamming Loss. This phase of training is the most optimal one, and the importance of chemical and biological modalities is determined.

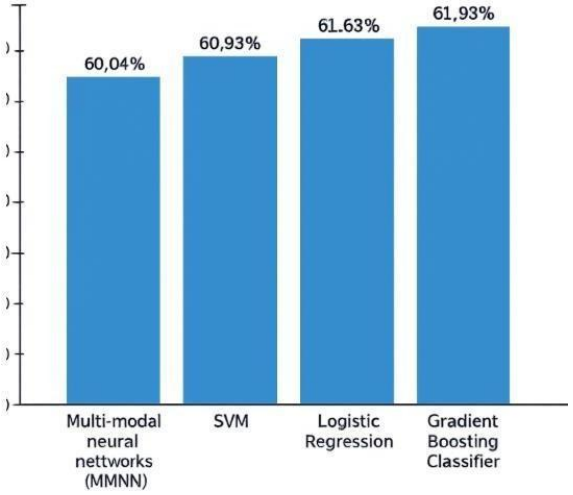

Fig.5. Final Output detection

Implementation will be finalized by deployment of the system which will be optimized and evaluated to ensure that the Deepside framework operates in the real life implementation. When a successful architecture has been found, e.g. the SMILESConv or multi-modal fusion model, compression mechanisms, such as pruning, quantization, and mixed-precision training are applied to reduce the computational intensity during inference. The model is released as formats that are consumer friendly such as ONNX or TorchScript that can be broadly compatible with cloud and edge computing hosts. A RESTful interface API is developed using either Flask or Django that can be utilized in the submission of new drug molecules either in SMILES sequence or chemical fingerprint format and receives an estimated score of ADR probability. There is the inclusion of visualization dashboard to present metrics, ROC curves and rankings of predictions of the ADRs, in an interactive manner. There is a lot of testing performed on withheld test sets and literature-based case studies to validate the accuracy of the prediction as well as to find the association with side-effects that are biologically plausible but not reflected in

existing datasets. The incorporation of automated prediction, interpretability, and flexible deployment services into the resulting system will turn the DeepSide research model into a valuable solution to drug discovery pipelines, toxicity screening systems, and pharmacovigilance systems to improve patient safety and reduce the risk of developing a drug.

CONCLUSION

DeepSide framework demonstrates the efficacy of the mixture of multi-modal biomedical information and advanced deep learning frameworks to forecast adverse drug reactions appropriately. The system, by incorporating the chemical structure properties, gene expression, and SMILES-based molecular representations, removes the flaws of the classical feature- engineered and similarity-based models. The results of the experiments show significant gains in macro- AUC and micro-AUC that attest to the high predictive capabilities of such models as MMNN and SMILESConv. The utilized framework can recognize common and common side effects and even does more discovery of possible ADRs not covered in existing datasets, which are worthy in the evaluation of early drug safety. DeepSide is more precise, scalable and generalizable to automated learning, reduced reliance on manual feature extraction and with the whole LINCS dataset. Overall, this research could be regarded as developing safer drugs, pharmacovigilance, and a foundation of future AI-based precision medicine.

REFERENCES

[1] Atias, N., & Sharan, R. (2011). A predictive algorithm of drug side effects. JCB published September 2004.

[2] Huang, C. Y., et al. (2011). Drug safety evaluation Systems pharmacology. Molecular BioSystems.

[3] Liu, M., et al. (2012). Nanoscale forecasting of drug side effects. Journal of the American Medical Informatics Association (JAMIA).

[4] Mizutani, S., et al. (2012). Mapping drug- protein interactions networks onto side effects. Bioinformatics.

[5] Yamanishi, Y., et al. (2012). Interpretation of drugs side-effects in terms of chemical and biological space. Bioinformatics.

[6] Shaked, B., et al. (2016). Drug side effects prediction based on metabolic networks. PLoS Computational Biology.

[7] Zielinski, M., et al. (2015). Side-effect pathogenesis and metabolic dysregulation. Cell Reports.

[8] Wang, Z., Clark, N., & Ma'ayan, A. (2016). Identification of drug side effects by

predicting with LINCS L1000 signatures. bioRxiv.

[9] Dey, H., et al. (2018). Prediction of adverse drug reaction using interpretable deep learning. Artificial Intelligence in Healthcare.

[10] Wang, L., et al. (2019). Deep learning to predict adverse drug events on clinical data. JAMIA.

[11] Wishart, D. S., et al. (2018). DrugBank: A complete drug discovery resource. Nucleic Acids Research.

[12] LINCS Consortium. (2017). The Library of Integrated Network-based Cellular Signatures (LINCS) Project. NIH.

[13] Tatonetti, N. P., et al. (2012). Inferring drug interactions using adverse-event reports. Science Translational Medicine.

[14] Krizhevsky, A., et al. (2012). Deep convolutional neural network imageNet classification. NIPS.

[15] Goodfellow, I., et al. (2016). Deep Learning. MIT Press.

[16] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature.

[17] Xu, K., et al. (2017). Molecular representation deep learning. Journal of Chemical Information and Modeling.

[18] Zhang, J., et al. (2019). Neural networks with molecular property prediction using SMILES. ACS Publications.

[19] Mikolov, T., et al. (2013). Effective approximate calculation of word representations in a vector space. ICLR.

U.S. Food and Drug Administration (FDA). (2020). Guideline on adverse drug reaction classification.