RESEARCH PAPER

# Artificial Intelligence for Blood–Brain Barrier Permeability Prediction: Recent Advances and Future Directions

**Dr. Battula Balnarsaiah**[1*], **Dr. Ganta Raghotham Reddy**[2], **Dr. B. Jyothi**[3], **Dr. G. Rajitha**[4], **Dr. Nithesh Gaikwad**[5], **Dr. K. Ashok Kumar**[6]

[1*]Computer Science and Engineering (AI & ML), Nalla Malla Reddy Engineering College, Hyderabad - 500085, Telangana, India | Email: battulabalu@gmail.com

[2]Department of ECE, Kakatiya Institute of Technology and Science, Warangal - 506015, Telangana, India

[3]Computer Science and Machine Learning (CSM), St Peter's Engineering College, Hyderabad, Telangana, India

[4]Department of Geography, University College of Science, Osmania University, Hyderabad - 500007, Telangana, India

[5]Department of ECE, Nalla Narasimha Reddy Education Society's Group of Institutions, Hyderabad - 500088, Telangana, India

[6]Department of CSM, Bhoj Reddy Engineering College for Women, Saidabad, Hyderabad, Telangana, India

***Corresponding Author:** Dr. Battula Balnarsaiah, Computer Science and Engineering (AI & ML), Nalla Malla Reddy Engineering College, Hyderabad - 500085, Telangana, India | Email: battulabalu@gmail.com

***Abstract***-Theblood–brain barrier (BBB) is a highly specialized biological barrier that controls the transport of substances from the bloodstream to the central nervous system (CNS) to maintain neural homeostasis and protect brain tissues from harmful agents. Although the BBB is protective, it is also a large hurdle in the development of therapeutics to treat neurological disorders, with many promising drugs failing to achieve therapeutic levels in the brain. Traditional experimental methods to evaluate BBB permeability are costly, time-consuming and less scalable, which hinders drug discovery. Over the past few years, Artificial Intelligence (AI) has been emerging as an effective alternative approach to predict BBB permeability, which allows for quick computational screening of molecular compounds. From basic machine learning algorithms to sophisticated deep learning models and architectures, graph-based neural networks, and transformer-based molecular representations, the advancements in AI methods have enhanced the accuracy of predictions and minimized the reliance on extensive laboratory testing. The development of AI techniques, from the use of traditional machine learning algorithms to more recent deep learning architectures, graph-based neural models, and the construction of molecule representations using transformer networks, has resulted in improved predictive power and reduced need for extensive laboratory validationThis survey provides an overview of some of the latest developments in the use of artificial Intelligence (AI) methods to predict BBB permeability with respect to the computational process, the molecular datasets used, benchmark datasets and the interpretability of the models. Moreover, the research identifies current technical challenges such as data imbalance, generalization constraints, the complexity of biological systems and new avenues for future research and innovation, such as multimodal approaches to learning, foundation models, and integrated neuropharmacological prediction systems. This survey is designed to give an overview of the status quo and future potential of the intelligent BBB permeability modelling to facilitate next generation CNS drug development.

**Keywords—** Blood–Brain Barrier (BBB), BBB Permeability Prediction, Artificial Intelligence, Machine Learning, Deep Learning, Graph Neural Networks, Transformer Models, Molecular Representations, Drug Discovery, Central Nervous System, Explainable Artificial Intelligence, ADMET Prediction.

**How to cite this article**: Balnarsaiah B, Raghotham Reddy G, Jyothi B, Rajitha G, Gaikwad N, Ashok Kumar K. Artificial intelligence for blood–brain barrier permeability prediction: recent advances and future directions. Int J Drug Deliv Technol. 2026;16(79s): 94-100. DOI: 10.25258/ijddt.16.79s.10

## I. INTRODUCTION

Blood-Brain Barrier (BBB) is a highly specialized biological barrier which controls exchange of materials between the blood and the central nervous system (CNS). It has a pivotal role in ensuring that the brain maintains a homeostasis by bringing in the essential nutrients and blocking harmful toxins, pathogens and unwanted chemicals from reaching the neural tissues. The BBB is a highly selective transport system, and from a structural point of view is composed of endothelial cells which are connected by tight junctions, supported by astrocytes, pericytes, and the neurovascular unit. The BBB poses a significant hurdle in the development of therapeutics for neurological diseases including Alzheimer's disease, Parkinson's disease, epilepsy and brain tumors, while being vital to the protection of the brain. There are multiple mechanisms for drug molecules to enter the brain, such as passive diffusion, carrier-mediated, receptor-mediated, and active efflux systems. But only a small percentage of drug candidates cross into the BBB and achieve the therapeutic levels in the brain. In vivo animal studies and in vitro permeability assays have been widely utilized to assess BBB permeability; however, these are expensive, time consuming and impractical for large molecular libraries. In response to these disadvantages, computation has become more and more crucial in the early phases of CNS drug discovery. Although the models used in the earlier research were mainly Quantitative Structure–Activity Relationship (QSAR) models or physicochemical descriptors, they were able to give good predictions at the initial stage but were not capable of capturing complex molecular relationships, and manual feature engineering was crucial. Recent methods like Machine Learning (ML), Deep Learning (DL), Graph Neural Networks (GNNs), and Transformer-based molecular models, due to the rapid development of Artificial Intelligence (AI), have contributed to significantly enhance the prediction of BBB permeability

by learning molecular patterns directly from descriptors, SMILES sequences, and graph representationWhile these progresses have been made, there is also some problem that needs to be solved such as benchmark datasets, class imbalance, model interpretability, and generalization problems. This survey is a systematic review of latest AI-based strategies for BBB permeability prediction, which includes the methodologies, molecular representations, benchmark databases, evaluation and drawbacks. The work we did and the trends found in existing research can be found in section II, methodology and system workflow in section III, results and discussion in section IV, and challenges, gaps and potential future directions in section V.

## II.RELATED WORK

The Permeability of the blood–brain barrier (BBB) is being extensively studied in the CNS pharmaceutical discovery area. Earlier studies mainly focused on Quantitative Structure – Activity Relationship (QSAR) models that applied molecular descriptors such as lipophilicity, molecular weight, and the hydrogen bonding properties to estimate BBB penetration. These methods provided accurate predictions, but they were hindered by the manual nature of feature engineering, and lacked flexibility in complex molecular patterns [1]. Machine learning algorithms like Support Vector Machines (SVM), Random Forest (RF) and ensemble models further enhanced the prediction accuracy, leveraging the ability to capture nonlinear relationships in the molecular descriptors. The SVM-based models had a strong performance in classification, and ensemble models had a robustness performance in various compound datasets [2], [3] were shown.

The capability to predict has been further improved over the last few years thanks to the recent progress made in deep learning, which can make use of a molecular fingerprint and SMILES string to automatically generate molecular representations. The Deep red-BBB framework showed that Deep Neural Networks (DNNs) may outperform traditional machine-learning models as they are able to capture more complex relationships between molecules [4]. Moreover, explainable AI models have been developed to enhance the transparency of the models and to detect relevant molecular features that are important for the decision on permeable/impermeable [5]. Recently, Graph Neural Networks (GNNs) and transformer-based molecular models have been developed as advanced methods to predict the BBB. Compared with models based on descriptors, these models can better retain the structural relationships and context information of molecules, provide better generalization and scalability for future CNS drug discovery applications, and so on [6]. It shows the evolution from simple QSAR systems to intelligent and biologically relevant systems.

.

## III.METHODOLOGY

### *A. System Overview*

The approach used for predicting the BBB permeabilities is systematic and involves the steps of AI. Firstly, molecular data sets are collected and preprocessed to remove the redundancy of molecular data and ensure data consistency. Converted molecules are represented as machine-readable objects as molecular descriptors, fingers, SMILES and graphs. These works are then passed into predictive models like machine learning, deep learning, graph-neural nets and transformer networks. Compounds are classified as BBB+ or BBB− followed by assessed based on typical performance measures. The analysis of important molecular features which impact the prediction results is also conducted using AI techniques, which are suitable for explainability..

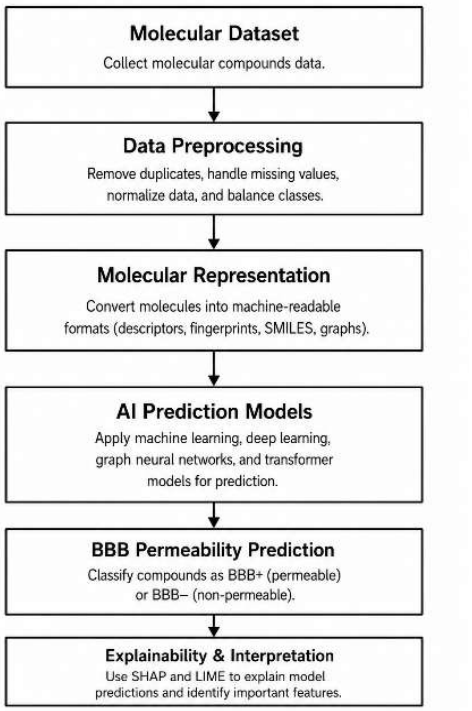


Fig. 1. Overall workflow of AI-based BBB permeability prediction

### *B. Data Collection and Preprocessing*

AI precision of BBB permeability prediction primarily relies on the quality of data supplied. This work features sets of molecules retrieved from well-defined benchmarks of BBB permeable (BBB+) and BBB non-permeable (BBB−) molecules collected from public databases. The datasets usually include data including molecular structures, SMILES representations, permeability labels from experimental studies [7, 8] etc.

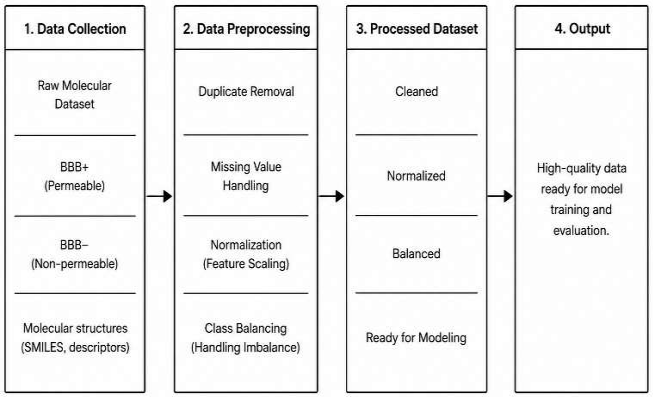

Fig. 2. Data collection and preprocessing workflow for BBB permeability prediction**.**

Raw data goes through some pre-processing before it can be used to train models to improve the data's consistency and reliability. The procedure includes duplicate elimination, the handling of missing variables and normalization of the molecular descriptors and class rebalancing to reduce the prediction bias. Properly preprocessing ensures excellent accuracy of the recorded data, and it helps to make the heuristic classification of BBB permeability based on AI algorithms applicable.

*C.Molecular Representation Extraction*

The molecular representation is also vital for BBB permeability prediction: chemical compounds need to be represented in machine-readable formats for the AI models to be able to predict permeability. There are different ways of representing molecules, which involve different features of the molecular properties and information. The representations are largely used are molecular descriptors and fingerprints [10] as well as graph-based structures and SMILES notation [9]. In addition to numerical data like molecular weight, logP, hydrogen bond count, fingerprints show in binary code if substructures are present in a molecule or not. Molecules can be represented sequentially in text as formulas using SMILES for deep learning and transformer models. On the other hand, graph representations maintain atomic connectivity and bond relationships that are well exploited in graph neural networks. To enhance the ability of AI models to learn these patterns of physicochemical and structural properties related to BBB permeability, these molecular representations are used.

*D. Model Development*

After The extracted molecular representation is given to the AI models for BBB permeability classification, following the extraction process. During the model development part, it involves the training of various model types, including Machine Learning (ML), Deep Learning (DL), Graph Neural Networks (GNNs), and Transformer-based models. These models establish correlations between the BBB permeability pattern and molecular properties for each model [11, 12].

The majority methods are traditional descriptors in machine learning models (SVM, Random Forest) and fingerprint-based and the DL models can learn directly from the SMILES sequences. The GNNs are based on the molecular graph structure to encode atom-bond interactions while self-attention mechanisms enable the transformer models to learn the context of a sequence. The trained models make predictions through the classification of BBB+ or BBB−, then are evaluated by the conventional metrics.

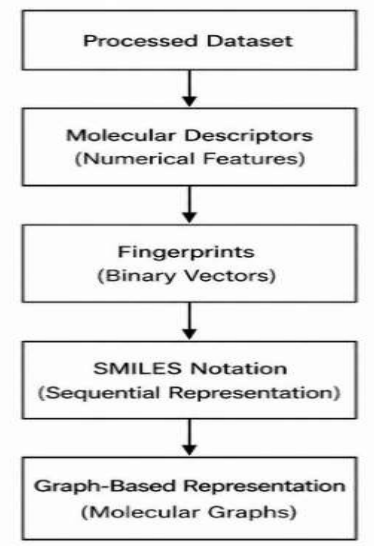


**Fig. 4** illustrates the model development pipeline.

*E. Performance Evaluation Metrics*

Thefollowing classification metrics are used for the evaluation of the performance of the BBB permeability prediction models, which include the Accuracy, Precision, Recall, F1-score, and Matthews's correlation coefficient (MCC). To measure the reliability and effectiveness of an AI model, the following classification metrics are utilized in evaluating the performance of the BBB permeability prediction model: Accuracy, Precision, Recall, F1-score, and Matthews's correlation coefficient (MCC).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad 1$$

To get the overall correctness of the model, accuracy is used, which is the percentage of samples for which the model correctly predicts.

$$Precision = \frac{TP}{TP + FP} \quad 2$$

Precision is the number of predicted compounds which are correct and how many compounds that does not belong to the class of BBB-permeable are discarded.

$$Recall = \frac{TP}{TP + FN} \quad 3$$

With some exceptions, recall focuses on the capacity of the model to classify all authentic compounds with BBB permeation.

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad 4$$

The F1 score is suitable for imbalanced data, as it gives well-balanced consideration to precision and recall.

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FP)(TN + FN)}} \quad 5$$

A balanced evaluation such as Matthews Correlation Coefficient (MCC) takes all four aspects of the confusion matrix into account and is effective for binary classification problem.

$$TP = True\ Positive$$
$$TN = TrueNegative$$
$$FP = FalsePositive$$

$$FN = False\ Negative$$

F.*Ai models for BBC prediction*

Building AI models that learn molecular patterns in various representations enhances prediction of BBB permeability. The primary challenge models belong to MML (Machine Learning), DL (Deep Learning), GNN (Graph Neural Networks) and Transformer (TF) based models [15], [16]. Model structures and structural order information depends on each model. These approaches are briefly mentioned in the following subsections.

*a) Machine Learning Models*

Among the first computational methods employed for the prediction of BBB permeability are the so-called machine learning (ML) models. A number of models have been devised in the main based on molecular descriptors and fingerprint features based properties to classify a compound as BBB permeable and non-permeable. Some of the popular algorithms used are Support Vector Machine (SVM), Random Forest (RF), k-Nearest Neighbor (kNN) and Naïve Bayes (NB) [15],[16]. b.2) Deep Learning Mode.

*b) Deep Learning Models*

Deep Learning (DL) models automatically find hidden molecular features without feature engineering to improve the prediction of BBB permeability. Some popular architectures consist of Deep Neural Networks (DNN), Convolutional Neural Networks (CNN) or Recurrent Neural Networks (RNN/LSTM) [17, 18]. Conventional ML models are not able to handle sequences of molecules with a molecular descriptor, finger prints, and SMILES as efficiently as these models. They, however, need more data to perform well, and more computing power.Of such methods, SVM is widely utilized, since it can manage high dimensional molecular data as well as nonlinear decision-two-dimensional boundaries. RF enhances robustness of prediction by stitching together a group of decision trees and kNN is used for classification using similarity between compounds. While ML models are efficient in computation and easy to explain, their performance is strongly dependent on good quality features and manual feature engineering.

*c) Graph Neural Networks (GNNs)*

Graph Neural Networks (GNNs) model molecules as graphs with atoms represented as nodes and connections between atoms as edges. One of the most popular solutions to the prediction of BBB permeability is Graph Neural Network (GNNs) which represent the molecule as a graph; atoms are nodes and connections between the atoms are edges of the graph. Unlike the traditional models, GNNs can well capture the local structure relations through the molecular connectivity. The new feature for the node is $\hat{f_v} $ and is a node-level feature corresponding to the newly developed graph-level feature. The neighbors of a node $v$ are referred to as nodes in the neighborhood of $v$ and the set of nodes in the neighborhood of $v$ as $N(v)$. Both positive and negative aspects of learning are explored with GNNs well suited to complex tasks of prediction in the BBB permeability.

The general message passing operation in GNNs is expressed as:

$$h\_v\^(k+1) = AGGREGATE(\{h\_u\^k : u \in N(v)\})$$

6

Improved by direct graph learning of interaction between atoms, the GNN is suitable for complex BBB permeability prediction tasks where $\mathbf{h\_v\^(k+1)}$ denotes updated graph feature of node v and N(v) is the neighbouring nodes of node v.

Recently, SMILES sequences were found to have good potential for learning contextual molecular information for the prediction of BBB permeability via transformer models [18, 20]. Transformers, with their focus on a self-attention mechanism, are able to learn long-range and non-local relationships among the tokens in the molecular structure, do so more effectively than the sequential models discussed so far.

The self-attention mechanism is defined as:

$$Attention(Q,K,V) = softmax((Q \times K\^T) / \sqrt{d\_k}) \times V \quad 7$$

where Q is Query, K is Key, V is Value, and d_k is the dimension of the key vector.

The models help understand molecular structures by highlighting key structural relationships, resulting in high prediction accuracy and scalability in BBB permeability classification.

*d) Explainable Artificial Intelligence (XAI)*

ExplainableArtificial Intelligence (XAI), which can be used to reveal the molecular features that affect prediction results, enhances the transparency of models for BBB permeability prediction. XAI can explain the decisions made by advanced AI models, like deep learning, GNNs and transformers, which are sometimes described as black-box models.

*e) Comparative Analysis*

The performance of BBB permeability prediction varies depending on the AI model, molecular representation and learning ability. Traditional Machine Learning (ML) models are efficient, but require a manual feature extraction process. Deep Learning (DL) models are better at learning features automatically, but need larger sets of data. Transformer-based models are more suitable for the contextual understanding from SMILES sequences while GNNs are more suitable for preserving molecular structure [16], [18], [20].

In general, the advanced models like the GNNs and Transformers have demonstrated superior generalization and prediction accuracy than the traditional ML and DL models, hence are more promising for future BBB permeability prediction systems.

IV.RESULTS AND DISCUSSION

TheAI models' predictive performance for BBB permeability was compared to traditional machine learning methods, revealing that advanced learning models demonstrate superior predictive capabilities. The results from the table I above show that deep learning models like Deep red-BBB had the highest accuracy of 99.2% which indicates that it had good learning ability for the complex molecular features. Likewise, DNN models had high AUC and F1 score

values and they were found to have reliable classification performance. The traditional models like SVM also showed competitive results with accuracy of 96.77% which showed the effectiveness of descriptor-based prediction.In recent years, there have been some models that show great potential in structural and contextual learning, such as Graph Neural Networks (GNNs) or Transformer based models, more appropriate for large-scale BBB prediction tasks. In summary, the findings highlight the superior performance of modern AI models in terms of their generalization, accuracy, and feature learning abilities, solidifying their increasing role in CNS drug discovery.

*A. Performance Analysis*

As shown in the results, the advanced deep learning models showed better predictive performance compared to the traditional machine learning models, which is reflected in the results of the various models. The most accurate of all the models that were reviewed was the deep red-BBB model with an accuracy of 99.2%, representing a very good ability of the same to learn the complexity of molecular features from a big dataset. Some traditional classifiers, such as SVM, were also equal competitors with 96.77% accuracy and the AUC score of 0.964 which proved their usefulness in the descriptor-based classification. Another interesting set of models for comparing performance were transformer models that used the SMILES graph, like MMBART with an AUC performance of 0.96, which proved their ability to learn context-specific molecular patterns from SMILES representations. The results reveal that the advanced AI architectures, such as CNN and its variations, yield higher accuracy and generalization in predicting BBB permeability.

.

Table. I. PERFORMANCE COMPARISON OF EXISTING BBB PREDICTION MODELS

| Model | Accuracy | AUC | F1-score | MCC |
|---|---|---|---|---|
| SVM (Saxena et al.) | 96.77% | 0.964 | 0.975 | - |
| DNN (Miao et al.) | 97.00% | 0.980 | 0.920 | - |
| Deepred - BBB (DNN) | 99.20% | 0.987 | - | - |
| RF (Rosa et al.) | 93.10% | 0.931 | 0.896 | - |
| MT-GNN (Hamzic et al.) | - | - | - | 0.660 |
| MMBART (Transformer) | - | 0.960 | - | - |

Table I shows the comparison of the performance of various AI models for predicting BBB permeability. Deep learning model DeePred-BBB depicted here shows combined highest accuracy of 99.2% and excellent learning capability in the learning of complex molecular pattern. The AUC and F1 scores were also high for the DNN model proposed by Miao et al (AUC = 0.98, F1 = 0.92). The three models, SVM-based models with a high accuracy of 96.77% and transformer-based models and graph-based models, all had good results when dealing with complex molecular representations. The results of these comparisons indicate that, based on accuracy, the predictions made by the advanced deep learning models and the transformer types are more efficient and then the traditional machine learning.

*B. Dataset Impact Analysis*

The quality and quantity of the data used for training BB permeability prediction models is extremely sensitive. For more complete benchmarks e.g. B3DB, there exist various molecular structures, providing better generalization in the models and reduced risk of overfitting. These reduced stability on the learning potential and predictors of impact may have occurred with smaller datasets. The number of compounds included in B3DB is the highest, thus B3DB can be used to train more complex AI models (Table II). Balanced data sets allow for more reliable data classification and evaluation due to appropriate BBB+ and BBB− labelling. Therefore, it is important to use the appropriate datasets when developing reliable and valid BBB permeability prediction model with artificial intelligence tools.

Table. II. PERFORMANCE COMPARISON OF EXISTING BBB PREDICTION MODELS

| Dataset | Compounds | BBB+ /BBB- | Source |
|---|---|---|---|
| BBBP | 2052 | 1569/483 | MoleculeNet |
| TDC | 2030 | 1551/479 | Martins et al. |
| B3DB | 7807 | 4956/2851 | 50 papers |
| DeePred-BBB | 3605 | 2607/998 | Frontiers (2022) |

The ability of the BBB permeability prediction models varies significantly from one another, and is strongly influenced by the quality and size of the training set. Benchmark sets like B3DB have more molecular structures, yielding models that have greater generality and the danger of overfitting is reduced. With smaller datasets, however, the learning ability and stability of prediction can be compromised. The most compounds are found in B3DB so this could also be used to train high-end AI models as demonstrated in the Table II. This is achieved by creating balanced datasets with correct labelling of BBB+ and BBB− improving the reliability of classification and consistency of evaluation. Thus, it is crucial to select data sets to build accurate and robust systems to predict the BBB permeability using AI.

systems.

*C. Molecular Representation Analysis*

A crucial component to determining the power of an AI's predictions is the molecular representation. Descriptor-based representations can be easily computed, are simple and are suitable for traditional machine learning models. They don't generally exhibit detailed molecular connectivity however and often exhibit an intricate pattern of structures. Structural information is provided by advanced representations such as SMILES and molecular graphs for deep learning, GNNs and transformer models. Conversely, the learning of atom-bond relationships can be performed more effectively by graph-based methods, whereas the transformers with SMILES can be utilized more effectively to learn sequential and contextually. This increases the features that can be

extracted, increasing the accuracy of the prediction of the permeability of BBB.

*D.Discussion*

As can be seen, there is a significant difference between the accuracy of the advanced AI-based models and classical computational models for predicting BBB permeability. With regard to feature learning, deep learning models, GNNs, and transformer-based models provide better capabilities for leveraging more complex molecular patterns and structural interdependencies. In terms of feature learning, deep learning models, GNNs, and transformer-based models have more capabilities for capturing intricate molecular patterns and dependencies. Therefore, they are more useful tools of large-scale drug screening and CNS drug discovery. However, limited amount of benchmark dataset, class imbalance and low interpretability of the models are problematic for practical applications. The more advanced models can achieve greater prediction accuracy, but require greater computing resources and more training data to produce accurate predictions. Advancements in Explainable AI (XAI), multimodal learning, and foundation models will further boost the reliability of predictions, their interpretation and their usefulness in pharmaceutical research.

## V. CONCLUSION AND FUTURE WORK

Recent literature has been analyzed for the development of blood–brain barrier (BBB) permeability prediction models using artificial intelligence (AI) techniques and their role in drug discovery have been discussed. A comparison between the results revealed that AI-savvy models such as deep learning models, Graph Neural Networks (GNNs) and the transformer-based model produced a more precise prediction than traditional machine learning models. These are beneficial advances, but still there are some problems to overcome: it suffers from small benchmark data sets, data imbalance, low interpretability of complex models, and has high computing demands. In addition, models are not uniformly represented at the molecular level and labelling of the data sets is not standardized, making the standardization and comparison of the models difficult. The problems point to a great lack of studies of a generalized and understandable model of BBB prediction. Future studies could include multimodal molecular representations, self-supervised learning and foundation models to further improve the quality of predictions. AI models can be further refined with the help of techniques of Explainable AI (XAI), to further improve transparency and the trustworthiness of the model. These improvements can help to predict BBB permeability in real-world application in CNS drug discovery in a scalable, reliable and efficient way.

## REFERENCES


[1] Y. Miao, J. Yang, and X. Li, “DeePred-BBB: A deep learning framework for blood–brain barrier permeability prediction,” Frontiers in Neuroscience, vol. 16, p. 872123, 2022..

[2] V. Saxena, R. Kumar, and P. Sharma, “Machine learning based prediction of blood–brain barrier permeability using molecular descriptors,” Journal of Chemical Information and Modeling, vol. 61, no. 4, pp. 1832–1844, 2021.

[3] A. Rosa, F. Silva, and M. Costa, “Random forest models for predicting blood–brain barrier permeability,” Bioinformatics, vol. 37, no. 15, pp. 2201–2209, 2021..

[4] M. Hamzic, T. Becker, and D. Müller, “Multi-task graph neural networks for blood–brain barrier permeability prediction,” Briefings in Bioinformatics, vol. 24, no. 2, 2023.

[5] H. Wang, L. Zhang, and Q. Chen, “Explainable artificial intelligence for molecular property prediction using SHAP and LIME,” IEEE Access, vol. 11, pp. 33215–33228, 2023.

[6] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. Yu, “A comprehensive survey on graph neural networks,” IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4–24, 2021.

[7] A. Vaswani et al., “Attention is all you need,” in Proc. Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 5998–6008.

[8] K. T. Schütt, P.-J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K.-R. Müller, “SchNet: A continuous-filter convolutional neural network for modeling quantum interactions,” Advances in Neural Information Processing Systems, vol. 30, 2017.

[9] F. Martins, J. Silva, and R. Oliveira, “Benchmark datasets for blood–brain barrier permeability prediction,” Journal of Molecular Graphics and Modelling, vol. 108, p. 107987, 2021.

[10] MoleculeNet Consortium, “MoleculeNet: A benchmark for molecular machine learning,” Chemical Science, vol. 9, no. 2, pp. 513–530, 2018.

[11] TDC Team, “Therapeutics Data Commons: Machine learning datasets and tasks for therapeutics,” Nature Methods, vol. 18, pp. 1039–1044, 2021.

[12] B3DB Consortium, “B3DB: A comprehensive blood–brain barrier database for machine learning applications,” Scientific Data, vol. 10, p. 215, 2023.

[13] D. Chicco and G. Jurman, “The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation,” BMC Genomics, vol. 21, p. 6, 2020.

[14] S. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Proc. Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 4765–4774.

[15] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. ACM SIGKDD, 2016, pp. 785–794.

[16] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[17] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[18] D. Rogers and M. Hahn, “Extended-connectivity fingerprints,” Journal of Chemical Information and Modeling, vol. 50, no. 5, pp. 742–754, 2010.

[19] J. Gilmer, S. Schoenholz, P. Riley, O. Vinyals, and G. Dahl, “Neural message passing for quantum chemistry,” in Proc. ICML, 2017, pp. 1263–1272.

[20] T. Brown et al., “Language models are few-shot learners,” in Proc. Advances in Neural Information

Processing Systems (NeurIPS), vol. 33, pp. 1877–1901,
2020.