predictBBB

predictBBB.ai

PREDICTING A COMPOUND'S BLOOD-BRAIN-BARRIER PERMEABILITY WITH LANTERN PHARMA'S AI AND ML PLATFORM, RADR®

Rick Fontenot[1], Panna Sharma[1]
[1] Lantern Pharma Inc., Dallas, TX, United States

Overview

This report presents methodologies and a series of machine learning generated models to predict the blood-brain-barrier (BBB) permeability for a collection of drugs in the Therapeutic Data Commons (TDC) BBB benchmark group challenge, originally published by Martins et al.. The semi-permeable BBB prevents the delivery of most drug compounds to the central nervous system (CNS) , limiting the efficacy of treatments in CNS disorders. The ability to predict which drugs are likely to pass through the BBB aids in identifying candidate treatments for disorders of the CNS. In this work, SMILES (Simplified Molecular Input Line Entry System) strings for drugs were transformed into molecular fingerprints and descriptor features along with already known important factors such as molecular weight and surface area for use in machine learning models with high accuracy in identifying a drug's ability to penetrate the BBB. After generating thousands of candidate features that could potentially be useful in predictions, different subsets of the most important features are selected for training models with logistic regression, random forest, SVM, and deep neural network methods as well as an ensemble of these base learner models. The final ensemble model is among the best in the field achieving a high accuracy of 94% with a sensitivity of 95% and specificity of 89% on the Therapeutic Data Commons test set.

Introduction

The blood-brain barrier serves to protect the central nervous system by tightly regulating the movement of substances between the blood and brain, thereby maintaining a stable, controlled environment essential for proper neural function. [1] The blood-brain-barrier (BBB) is a highly selective semi-permeable feature that serves to protect the central nervous system (CNS) and presents challenges in the effective delivery of therapeutic compounds for the treatment of brain and CNS disorders. Recent research indicates that only about 2—6% of small-molecule drugs can cross the blood—brain barrier (BBB), highlighting the significant challenge in delivering therapeutics to the central nervous system. [2] While the blood-brain barrier is a highly complex and selective interface, it allows substances to cross through several distinct pathways that collectively regulate brain access. Small lipophilic molecules cross the BBB by passive diffusion, essential nutrients such as glucose and amino acids are transported via specific carrier proteins, larger biomolecules like insulin and transferrin enter through receptor-mediated transcytosis (RMT), positively charged proteins and peptides may cross via adsorptive-mediated transcytosis (AMT), small hydrophilic molecules occasionally pass through paracellular routes under disrupted conditions, while efflux pumps actively remove many xenobiotics and lipophilic drugs back into the blood, limiting their CNS access [3]

When developing machine learning models to predict blood-brain barrier permeability, it is essential to account for the distinct physicochemical and transport characteristics of small versus large compounds, as these groups often utilize fundamentally different mechanisms to cross the barrier. For modeling passive diffusion across the blood-brain barrier (BBB), a molecular weight threshold of 500 daltons (Da) is commonly employed to define small molecules, as compounds exceeding this size are significantly less likely to permeate the BBB without active or facilitated transport mechanisms. [4]

Building machine learning models to predict passive diffusion across the blood-brain barrier (BBB) is more feasible due to the availability of extensive publicly accessible datasets with labeled outcomes for small molecules, such as the B3DB dataset comprising over 7,800 compounds with BBB permeability annotations. [5] In contrast, data for larger molecules that rely on active or facilitated transport mechanisms are scarce, less standardized, and often proprietary, making model training and validation significantly more challenging. Labels in the B3DB dataset were primarily derived from in vivo rodent studies using methods like logBB, brain uptake index, and microdialysis, which aim to capture passive diffusion but may inadvertently reflect active transport mechanisms as well.

While BBB permeability data from animal models come with limitations due to species-specific differences in transporters, metabolism, and dosing routes, they still offer valuable, biologically grounded insights that enable the development of predictive models, especially for passive diffusion. A study by Gülave et al. (2025) demonstrated that machine learning models trained on in vivo unbound brain-to-plasma partition coefficient (Kp,uu,BBB),BBB data can reliably predict human BBB transport when integrated into physiologically-based pharmacokinetic (PBPK) models, as the resulting simulations closely matched observed human CNS drug exposure. [6]

One limitation of machine learning models trained on in vivo data to predict blood—brain barrier (BBB) permeability is their inability to account for BBB disruption and leaky barrier states, which can significantly alter drug permeability in ways not reflected in standard experimental labels. Blood—brain barrier (BBB) disruption refers to the breakdown of the BBB's selective permeability, allowing normally restricted substances to enter the brain. This disruption can result from various factors, including inflammation, trauma, disease pathology, or drug-induced mechanisms that impair endothelial integrity or tight junction function.

Prior work utilizing machine learning models to predict BBB permeability were created by generating molecular fingerprints as features using software packages such as PaDEL-Descriptor software [7] or DeepChem and RDKit [8] and produced results with AUC-ROC ranging from 0.849 [7] to 0.905 [8] and accuracy of 0.798 [7].

While molecular fingerprints as machine learning features have shown this predictive success, previous research also suggests that BBB permeability decreases as the molecule's surface area increases [9]. BBB permeability also decreases as hydrogen bonds are added to the structure, whether it is a hydrogen bond donor or hydrogen bond acceptor [9]. Incorporating these features for drugs in addition to the fingerprints previously shown to be valuable may provide for higher prediction accuracy.

Although BBB permeability prediction with machine learning and deep learning methods based on molecular fingerprints have demonstrated accuracy beyond short rule based criteria, expanding the number of descriptors available as candidate features may improve accuracy. The workflow approach here expands the candidate features based on insights from literature reviews, then reduces the final amount of features used in training models based on importance in feature selection algorithms. In addition to better model performance on test data, these models use fewer features and are more robust and interpretable.

Interpretation of key features can improve the understanding of how specific chemical structures impact BBB permeability, and how molecule design can be improved to enhance permeability.

Methods

Data with small molecule BBB behavior labels used for training and testing were compiled from two sources. The Therapeutic Data Commons (TDC) [10] hosts a dataset that is derived from splits of compounds originally published by Martins et al [11]. TDC hosts a leaderboard for public model entries, and this test set represents a benchmark for direct comparison to competitive modeling approaches in the field. While the TDC test set is valuable for direct comparison to other models, increasing the number of compounds in the training set is critical to improving accuracy. A larger set of compounds for training was obtained from the B3DB data set [5]. The combination of these datasets after removing duplicates resulted in 9,902 compounds available, plus the 406 compounds already split into the TDC test set. These available compounds were further split into 7,273 for training, 909 for validation, and 910 for a second test set composed of compounds from the B3DB dataset for comparison of metrics to the TDC test set.

Simplified Molecular Input Line Entry System (SMILES) structures were used to translate a chemical's three-dimensional structure into a string of symbols that can be processed by computer software programs. The “Rdkit” python library was used to convert the SMILES drug structures into numerical features such as Morgan fingerprints , rdk fingerprints, MACCS fingerprint, and descriptors including 2D and 3D Autocorrelations as well as 3D Getaway and WHIM descriptors. [12] These binary and non-binary numerical features served as a proxy for different atomic properties including element connectivity, chemical features, bond type, atomic mass, and electrotopological state. Well established atomic rules (e.g. lipinski rules, ghose filter, veber filter etc.) and their attributes were also generated as features. These processes generated 4,939 candidate features for each drug (Table 1), which were subsequently fed into feature reduction methods prior to machine learning model generation.

Feature Generation Method	# of Features	Values
Rdk fingerprints	2048	Binary
Morgan fingerprints	2048	Binary
MACCS fingerprints	167	Binary
2D Autocorrelation descriptors	192	Continuous
3D Autocorrelation descriptors	80	Continuous
Rules/filters (6) and its attributes (11)	17	Mix
Getaway 3D descriptors	273	Continuous
WHIM 3D descriptors	114	Continuous
Total	**4939**	**Mix**

Table 1. Details of feature generation methods

Following the initial candidate feature generation, feature reduction was performed to reduce feature multicollinearity prior to feature selection and modeling. Tanimoto similarity was used to screen binary features such as fingerprints and Pearson's R was used to screen numeric features. For both metrics one feature was dropped in each pair of features with a correlation higher than 95% resulting in dropping 19 of the original binary features and 284 of the original continuous features.

Next chi-square tests were performed for all individual fingerprint features to determine whether permeability is dependent on the fingerprint. Fingerprints with a p-value less than 0.05 were considered as significant in the first phase. In the second phase, the ratio of permeable and non-permeable samples were calculated for drug samples containing each fingerprint. Fingerprints with a permeability of 50% or lower were categorized as significantly negatively associated fingerprints, and those with 80% or greater permeability were considered significantly positively associated fingerprints. Based on these results newly engineered features were created totaling the count of negatively and positively associated fingerprints for each drug sample.

Following feature generation, reduction and engineering steps described above, features were transformed using degree 2 poly kernel principal component analysis.

Due to the class imbalance between the majority of drug samples being permeable (65%) and a minority of samples being non-permeable (35%), Synthetic Minority Oversampling Technique (SMOTE) [13] was performed on the training set of data utilizing SVMSMOTE with 12 k-nearest neighbor algorithm to create synthetic data for the minority class until the sample counts were balanced between permeable and non-permeable observations. This augmentation of the training data set served as the input to feature selection and model training phases of the workflow. No augmented samples were generated for the validation or test sets of data.

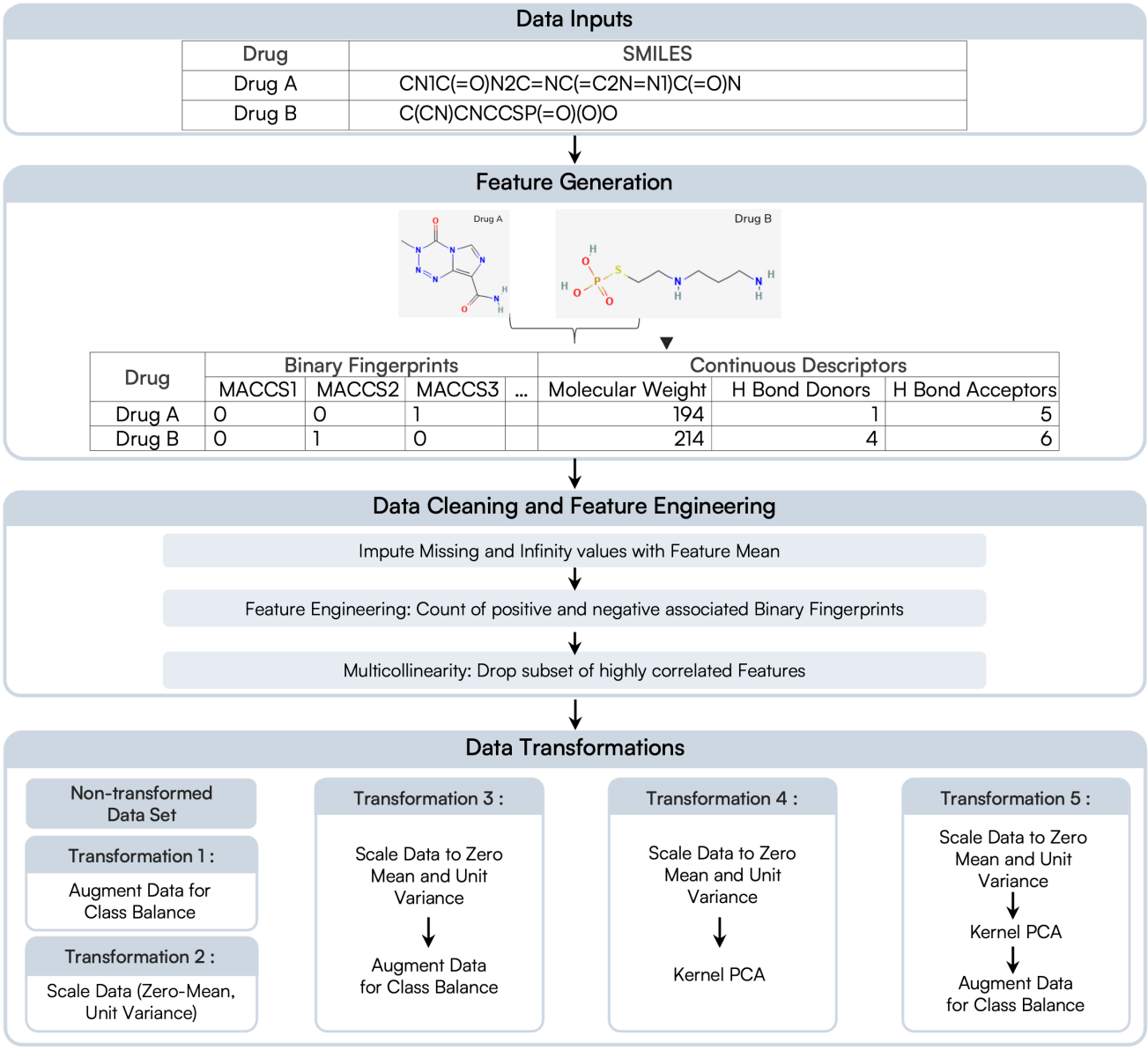

Figure 1: Data Preparation Workflow

Candidate features were reduced using logistic regression with the least absolute shrinkage and selection operator penalty to shrink the least important feature's coefficients to zero thus eliminating from features selected for use in model training. Ten-fold cross-validations were performed on a search range for the L1 regularization parameter value, which provided the highest average accuracy across the folds. The L1 regularization value identified as optimal in the search was used to train a final feature selection model that eliminated features with a coefficient of zero and ranked remaining features in order of importance by the absolute value of their coefficient. Feature selection lists were generated for the original un-transformed features as well as the kernel PCA transformed features to evaluate the use of each feature set on each predictive modeling type.

Four base learner methods of modeling were utilized across all methods of data transformation sets to generate diverse methods of predictions that serve as inputs to an ensemble meta-learner in a subsequent phase.

The first base learner method used was a logistic regression. Previous implementation during feature selection utilized L1 regularization to reduce features. The base learner model included a further search using ten-fold cross validation to determine the optimal L2 regularization parameter values. This model used the kernel pca transformed features without oversampling augmentation.

The second base learner method used was a Deep Neural Network. The design of the neural network was generated using a search on the optimal architecture for a range of two to five fully connected dense hidden layers. Each hidden dense layer included L2 regularization and was followed by a dropout layer. The search additionally included a range of neurons used in each hidden layer, a selection of optimizers, and parameters for learning rate reductions over the course of training. Each iteration of the architecture search included early stopping criteria using a holdout subset from the training data to stop model training at the epoch which represented the highest area under the receiver operating characteristic curve (AUC-ROC) on the holdout samples. A final model was constructed with the 4 hidden layers consisting of [67, 70, 84, 88] neurons with [relu, tanh, relu, relu] activation functions,, and dropout rates of [0.029, 0.122, 0.039, 0.162] and L2 regularization rate of 1.6e-8. The model used the RMSprop optimizer with an initial learning rate of 1e-3 which is reduced by a factor of 0.5 after 3 epochs without improvement in loss. identified in the search and was fully trained up to the early stopping criteria. This model included kernel pca transformed features without augmentation at the feature selection stage but with augmentation during training.

The third base learner method used was a random forest. This base learner model included a search using ten-fold cross-validation to determine the optimal number of estimators, depth, and minimum samples per split and leaf to reduce the effects of overfitting. This model included original features without kernel pca transformation and used augmentation for both the feature selection and model training processes.

The fourth modeling method was a Support Vector Machine (SVM). This base learner model included a search using ten-fold cross-validation to determine the optimal kernel, degree, regularization strength, and kernel gamma coefficient which resulted in a final model with a linear kernel and regularization strength equal to 2. This model included kernel pca transformed features without augmentation at the feature selection stage but with augmentation during training.

After training each of the base learner models, the predicted probability of permeability was calculated on holdout validation samples that were not included in the model's training samples. These validation sample predictions were subsequently used to train an ensemble meta-learner.

The fifth modeling method was an ensemble method where validation sample predictions from all base learner models were used as feature inputs to a logistic regression meta-learner ensemble model. All base models were evaluated as meta-learner inputs and permutations of the base-learner combinations. The combination of base-learners with the highest area under the receiver operating characteristic curve was selected as the final meta-learner.

Generating predictions on the test set for evaluation was completed in two phases. First, the probability of permeability for each sample was predicted for each base learner and its associated selected features. Second, the validation set base learner probabilities were used as inputs to train a logistic regression meta-learner ensemble model used to make the final predictions on the test set. BBB permeability classification labels were assigned using a threshold of greater than 0.5 predicted probability as the drug being permeable for all model types for the purposes of reporting accuracy, sensitivity and specificity scores.

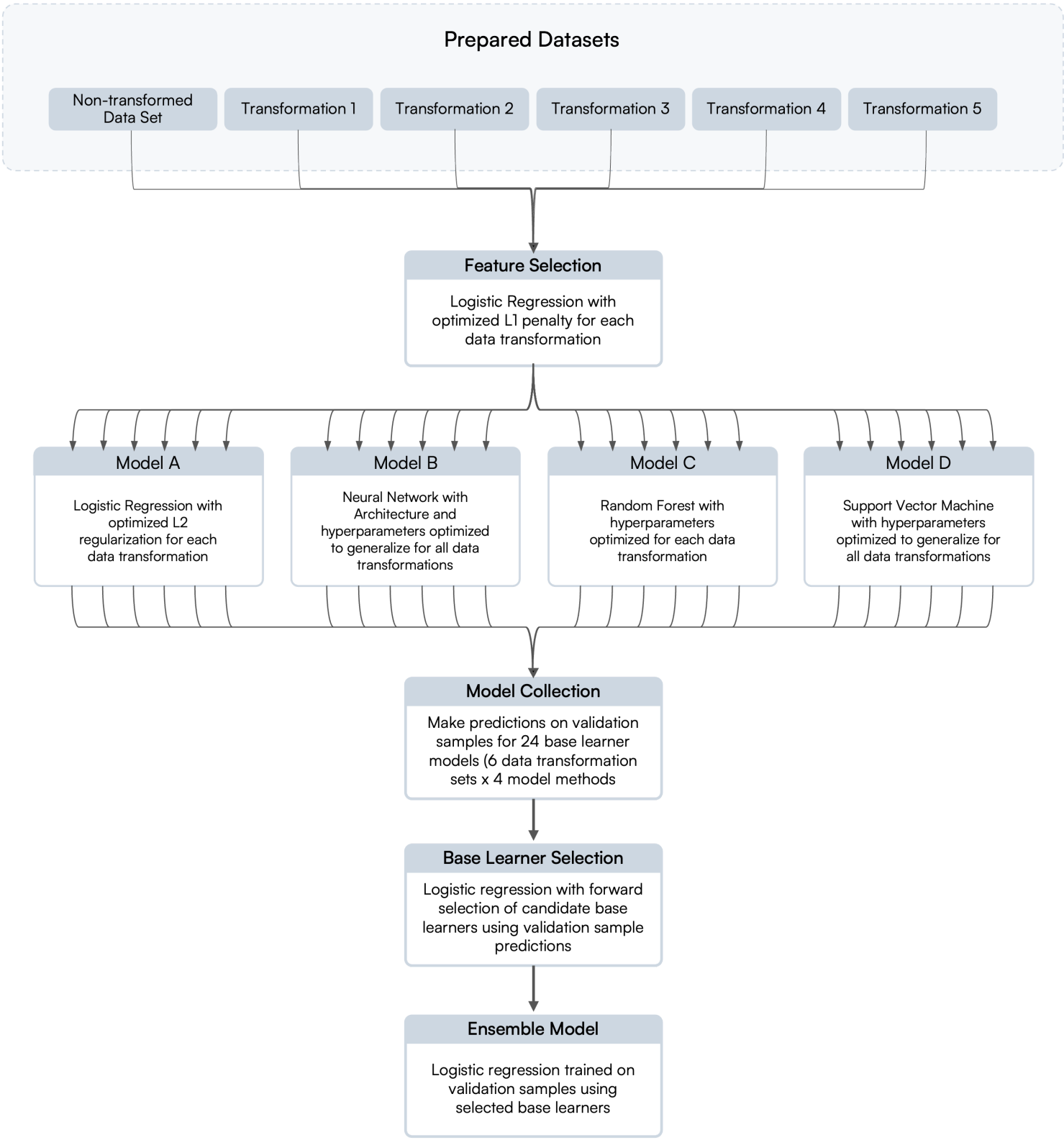

Figure 2: Modeling Workflow

While model tuning and training are time intensive, once the trained models are preserved molecular descriptor generation and BBB predictions on additional compounds is scalable. Utilizing cloud compute resources with 16 cores and 128 GB RAM, on average each compound in a large set of SMILES strings requires 3 seconds, with predictions on a batch of 100,000 compounds completed in 3.5 days.

Performance of the models were evaluated by the following metrics:

Area Under the Receiver Operating Characteristic Curve (AU-ROC): provides a single scalar value that summarizes the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR) across various decision thresholds.

True Positive Rate (TPR) calculated from True Positive (TP) and False Negatives (FN)

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

False Positive Rate (FPR) calculated from False Positives (FP) and True Negatives (TN)

$$FPR = \frac{FP}{FP + TN}$$

The ROC curve is a plot of the TPR against the FPR at different probability threshold levels. The area under this curve (AU-ROC) gives a scalar value that captures the model's ability to discriminate between the positive and negative classes.

Accuracy: measure of how often the classifier is correct overall

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

Sensitivity: measure of all the actual positive cases, how many did the model correctly detect. A higher sensitivity means the model is better at identifying positive cases

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

Specificity: measure of all the actual negative cases, how many did the model correctly identify. A higher specificity indicates that the model is less likely to give false positives

$$Specificity = \frac{TN}{TN + FP}$$

Results

The ensemble model obtained the highest balance of Accuracy, Sensitivity, Specificity and AU-ROC by blending the diverse predictions of the deep neural network, random forest, and SVM models as base learners. All four machine learning model types generated have AUC-ROC scores ranging from 0.92-0.96, which placed them at the top of the BBB-Martins leaderboard as follows as of the date of submission.

Specifically the ensemble model achieves 94.1% accuracy on the TDC test set of 406 unseen molecules. Expanding beyond with a new rigor of validation on an additional set of 910 unseen molecules split from the B3DB set, the AU-ROC is a similar 0.97 with accuracy of 90.4%. These results observed on two different test sets totalling 1,316 unseen molecules demonstrates the generalizability of the ensemble model.

						Data Preparation		
Model Type	Ensemble Weight	Accuracy	Sensitivity	Specificity	AU-ROC	Scaled	Data Augmentation	Kernel-PCA
Ensemble		**94.1%**	**95.4%**	**88.5%**	**0.964**			
Random Forest	3.353	93.3%	95.7%	83.3%	0.961		X	
Logistic Regression	1.705	89.7%	90.9%	84.6%	0.933		X	
Random Forest	1.182	92.9%	96.6%	76.9%	0.937	X	X	X
Neural Network	1.073	93.6%	95.7%	84.6%	0.965	X	X	X
Neural Network	0.687	93.6%	95.7%	84.6%	0.964	X		X
Logistic Regression	-0.131	88.4%	88.7%	87.2%	0.945	X		
Random Forest	-1.128	90.9%	98.8%	57.7%	0.904	X		X

Table 2. Model BBB Prediction Performance on TDC test set (406 drugs)

						Data Preparation		
Model Type	Ensemble Weight	Accuracy	Sensitivity	Specificity	AU-ROC	Scaled	Data Augmentation	Kernel-PCA
Ensemble		**90.4%**	**93.1%**	**85.6%**	**0.969**			
Random Forest	3.353	89.0%	91.9%	83.8%	0.966		X	
Logistic Regression	1.705	89.5%	90.8%	87.8%	0.942		X	
Random Forest	1.182	88.5%	94.6%	77.2%	0.956	X	X	X
Neural Network	1.073	91.0%	92.4%	88.4%	0.962	X	X	X
Neural Network	0.687	90.3%	93.9%	83.8%	0.963	X		X
Logistic Regression	-0.131	88.6%	89.0%	87.8%	0.947	X		
Random Forest	-1.128	85.9%	97.8%	64.0%	0.928	X		X

Table 3. Model BBB Prediction Performance on B3DB test set (910 drugs)

Limitations, Challenges and Future Directions

Improvements to Small Molecule Modeling

Efflux pumps at the blood—brain barrier (BBB) are specialized transporter proteins that actively expel a wide range of drugs and xenobiotics from the brain back into the bloodstream, serving as a key defense mechanism that limits central nervous system drug exposure. Efflux pumps are primarily represented by three members of the ATP-binding cassette (ABC) transporter family—P-glycoprotein (P-gp/ABCB1), breast cancer resistance protein (BCRP/ABCG2), and several multidrug resistance-associated proteins (MRP1/ABCC1 and MRP4/ABCC4)—that work together to actively export a broad spectrum of drugs and xenobiotics back into the bloodstream [14]. Permeability-glycoprotein (P-gp) substrates consistently show poor brain penetration due to active efflux at the blood—brain barrier (BBB), whereas P-gp inhibitors can increase CNS drug exposure by blocking this efflux, and non-substrates/non-inhibitors bypass P-gp entirely, relying solely on passive diffusion [15]

While the generated molecular features and current modeling methods could simultaneously model both the effects of passive diffusion for entry and the efflux pumps counteracting entry, the relevant features for these two impacts may be different and the weight and direction of correlations to behaviors may be different. One potential improvement to modeling small compound behavior at the BBB, could be first modeling inhibitor and substrate classifications of compounds for the most important ABC transporter proteins such as PGP and BCRP. These predictions could be used as additional features or as partitions into separate training sets for a subsequent BBB model.

Modeling Larger Molecules

Larger molecules could cross the BBB by two different mechanisms, adsorptive-mediated transcytosis (AMT) or receptor-mediated transcytosis (RMT).

Adsorptive-mediated transcytosis enables certain large or charged molecules to cross the blood—brain barrier by first binding electrostatically to the negatively charged endothelial surface, followed by uptake via clathrin- or caveolae-mediated vesicular transport and subsequent release on the brain side [16]. Adsorptive-mediated transcytosis (AMT) across the blood—brain barrier is driven by characteristics such as positive surface charge (e.g., high isoelectric point or zeta potential), molecular size and shape suitable for vesicular uptake, and peptide/protein sequence features like cationic motifs found in cell-penetrating peptides, which collectively facilitate electrostatic binding to the endothelial surface and endocytic transport [16]. Machine Learning models should be suited to learn from experimental BBB uptake data of AMT-reliant molecules (e.g., cationic peptides, cell-penetrating peptides).

BBB penetration via receptor-mediated transcytosis requires that therapeutic molecules engage specific endothelial receptors—such as transferrin or insulin receptors—with optimized binding affinity that promotes internalization and trafficking across the BBB without triggering receptor degradation or intracellular retention [17]. Since RMT relies on specific receptor-ligand interactions, molecular docking simulations to predict binding affinity, pose, and epitope accessibility could be necessary for understanding receptor engagement and initiating transcytosis. Since RMT success depends on more than binding—trafficking, release, receptor recycling, and endosomal sorting, a combination of docking and machine learning could be the most appropriate approach. However the need for human in the loop setup and interpretation of docking simulations could cause challenges in scalability for screening a large list of compounds for the ability to cross the BBB via RMT

Due to the lack of publicly available datasets with sufficient labeled compounds with respect to AMT and RMT behaviorOpen to collaborative efforts, training models are currently a challenge and may require collaboration with organizations in possession of large private datasets.

Collaboration Needs for BBB Research

Training effective models for AMT and RMT behavior prediction faces a significant challenge: the scarcity of publicly available datasets with adequately labeled compounds. To address this limitation, we are actively seeking collaborative partnerships with organizations that possess large private datasets relevant to blood-brain barrier research.

Collaboration Goals:

- Partner with pharmaceutical companies, research institutions, and biotech organizations that have proprietary compound databases with measured results on BBB behaviors. For compounds that cross the BBB, additional information on the mechanism by which the compound crosses would allow for separate AMT and RMT models.
- Collaborative development of models and machine learning approaches that aid in accurate, scalable modeling of active transport mechanisms
- Create standardized labeling protocols to ensure dataset compatibility across collaborators.

We welcome discussions with potential partners who share our commitment to advancing BBB research through shared resources and expertise.

References:

1. N. Joan Abbott, Adjanie A.K. Patabendige, Diana E.M. Dolman, Siti R. Yusof, David J. Begley, Structure and function of the blood—brain barrier,Neurobiology of Disease, Volume 37, Issue 1, 2010, Pages 13-25, ISSN 0969-9961, https://doi.org/10.1016/j.nbd.2009.07.030.
2. Fleur M.G. Cornelissen, Greta Markert, Ghislaine Deutsch, Maria Antonara, Noa Faaij, Imke Bartelink, David Noske, W. Peter Vandertop, Andreas Bender, Bart A. Westerman,
Explaining Blood—Brain Barrier Permeability of Small Molecules by Integrated Analysis of Different Transport Mechanisms, Journal of Medicinal Chemistry, Volume 66, Issue 11, 2023, Pages 7253-7267, ISSN 1520-4804, https://doi.org/10.1021/acs.jmedchem.2c01824.
3. Fu, B.M. (2018). Transport Across the Blood-Brain Barrier. In: Fu, B., Wright, N. (eds) Molecular, Cellular, and Tissue Engineering of the Vascular System. Advances in Experimental Medicine and Biology, vol 1097. Springer, Cham. https://doi.org/10.1007/978-3-319-96445-4_13
4. Harris, W.J., Asselin, MC., Hinz, R. et al. In vivo methods for imaging blood—brain barrier function and dysfunction. Eur J Nucl Med Mol Imaging 50, 1051—1083 (2023). https://doi.org/10.1007/s00259-022-05997-1
5. Meng F, Xi Y, Huang J, Ayers PW. A curated diverse molecular database of blood-brain barrier permeability with chemical descriptors. Sci Data. 2021 Oct 29;8(1):289. doi: 10.1038/s41597-021-01069-5. PMID: 34716354; PMCID: PMC8556334.
6. Gülave, B., van den Maagdenberg, H.W., van Boven, L. et al. Prediction of the Extent of Blood—Brain Barrier Transport Using Machine Learning and Integration into the LeiCNS-PK3.0 Model. Pharm Res 42, 281—289 (2025). https://doi.org/10.1007/s11095-025-03828-0
7. Liu L, Zhang L, Feng H, Li S, Liu M, Zhao J, et al. Prediction of the Blood—Brain Barrier (BBB) Permeability of Chemicals Based on Machine-Learning and Ensemble Methods. Chem Res Toxicol. 2021;34:1456—67.
8. Tian H, Ketkar R, Tao P. Accurate ADMET Prediction with XGBoost. Arxiv. 2022. https://doi.org/10.48550/arxiv.2204.07532.
9. Profaci CP, Munji RN, Pulido RS, Daneman R. The blood—brain barrier in health and disease: Important unanswered questions. J Exp Medicine. 2020;217:e20190062.
10. Huang, K., Fu, T., Gao, W., Zhao, Y., Roohani, Y., Leskovec, J., Coley, C. W., Xiao, C., Sun, J., and Zitnik, M. Therapeutics data commons: Machine learning datasets and tasks for drug discovery and development. arXiv preprintarxiv:2102.09548, 2021
11. Martins IF, Teixeira AL, Pinheiro L, Falcao AO. A Bayesian approach to in silico blood-brain barrier penetration modeling. J Chem Inf Model. 2012 Jun 25;52(6):1686-97. doi: 10.1021/ci300124c. Epub 2012 Jun 6. PMID: 22612593.
12. RDKit: Open-source cheminformatics. https://www.rdkit.org https://doi.org/10.5281/zenodo.7671152.
13. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. J Artif Intell Res. 2002;16:321—57.
14. Helms, H.C., Hersom, M., Kuhlmann, L.B. et al. An Electrically Tight In Vitro Blood—Brain Barrier Model Displays Net Brain-to-Blood Efflux of Substrates for the ABC Transporters, P-gp, Bcrp and Mrp-1. AAPS J 16, 1046—1055 (2014). https://doi.org/10.1208/s12248-014-9628-1
15. Na Li, Priyanka Kulkarni, Akshay Badrinarayanan, Adey Kefelegn, Raffi Manoukian, Xingwen Li, Bhagwat Prasad, Matthew Karasu, William J. McCarty, Charles G. Knutson, Anshul Gupta, P-glycoprotein Substrate Assessment in Drug Discovery: Application of Modeling to Bridge Differential Protein Expression Across In Vitro Tools, Journal of Pharmaceutical Sciences, Volume 110, Issue 1, 2021, Pages 325-337, ISSN 0022-3549, https://doi.org/10.1016/j.xphs.2020.09.017.
16. Hervé, F., Ghinea, N. and Scherrmann, J.M., 2008. CNS delivery via adsorptive transcytosis. AAPS Journal, 10(3), pp.455—472. https://doi.org/10.1208/s12248-008-9055-2
17. Haqqani, A.S., Bélanger, K. and Stanimirovic, D.B., 2024. Receptor-mediated transcytosis for brain delivery of therapeutics: receptor classes and criteria. Frontiers in Drug Delivery, 4. https://doi.org/10.3389/fddev.2024.1360302