Series on Biomechanics, Vol.39, No.1 (2025),74-84
DOI: 10.7546/SB.01.10.2025

Application of artificial intelligence and machine learning to study the structure–biological activity relationship

F. Sapundzhi[a], S. Georgiev[b,c,*], M. Lazarova[d,*], M. Popstoilov[a]

[a] *Dept. of Communication and Computer Engineering, Faculty of Engineering, Neofit Rilski South-West University, 66 Ivan Myhailov Str., 2700 Blagoevgrad, Bulgaria, email: sapundzhi@swu.bg*
[b] *Dept. of Informational Modeling, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 8 Acad. Georgi Bonchev Str., 1113 Sofia, Bulgaria, email: sggeorgiev@math.bas.bg*
[c] *Dept. of Applied Mathematics and Statistics, Faculty of Natural Sciences and Education, A. Kanchev University of Ruse, 8 Studentska Str., 7004 Ruse, Bulgaria, sggeorgiev@uni-ruse.bg*
[d] *Faculty of Applied Mathematics and Informatics, Technical University of Sofia, 8 Kliment Ohridski Blvd., 1000 Sofia, Bulgaria, meglena.laz@tu-sofia.bg*

Received: 16 December 2024 *Accepted: 19 March 2025* **Corresponding Author*

Abstract

Objective: This study aims to investigate the structure–biological activity relationship (SAR) of delta-opioid ligands using artificial intelligence and machine learning approaches, advancing previous efforts based on polynomial modeling. The goal is to establish nonlinear relationships between molecular docking results, binding energies, and biological activities, thereby improving predictive accuracy in drug design. **Materials and methods:** Three delta-opioid receptor (DOR) models were used: a theoretical model (PDB: 1ozc), a crystal structure model (PDB: 4ej4), and a homology-modeled structure. Docking studies were performed using software GOLD, and binding free energies were calculated using Molegro Molecular Viewer (MMV). Polynomial modeling was used as a baseline, while machine learning regression techniques, including k-Nearest Neighbors, Gradient Boosting, Random Forest, and Extremely Randomized Trees, were applied to capture nonlinear SARs. Model performance was evaluated using k-fold cross-validation and grid search optimization. **Results:** Among the tested models, Gradient Boosting exhibited the highest predictive accuracy, outperforming polynomial regression. Statistical metrics such as coefficient of determination R^2, root mean square error (RMSE), and sum of square errors (SSE) demonstrated the efficacy of the machine learning approach in capturing complex SARs across all three DOR models. **Discussion:** Machine learning algorithms provide a robust and efficient method for predicting SARs, enabling more precise identification of drug candidates. Compared to polynomial modeling, the proposed methods exhibit greater flexibility and reliability in uncovering nonlinear relationships. **Conclusion:** The integration of machine learning into SAR analysis enhances predictive capabilities, accelerating drug discovery and optimization. Future work will focus on extending these methods to other receptor-ligand systems and exploring additional algorithms.

Keywords: Structure–biological activity relationship, delta-opioid receptor, molecular docking, machine learning, drug design

1. Introduction

In the identification and development of drugs, computer-aided drug design (CADD) employs a diverse array of computational tools and methods [1]. The drug design process relies on computer modeling and docking experiments [2]. Using these techniques, we can predict and analyze the interactions between potential drug molecules and target biological structures, such as proteins. To illustrate how computer modeling, docking experiments, and machine learning contribute to drug design, here are a few key facts (Figure 1).

Structure-Based Drug Design (SBDD): We can study the three-dimensional (3D) structures of biological molecules, through computer modeling, especially targets involved in many diseases [3]. Our researchers will benefit from this information in designing drugs that modulate the activity of specific target

proteins. SBDD includes the following approaches: Homology modeling, Molecular dynamics, Molecular docking and scoring functions, Machine learning, etc. [4].

Homology modeling (HM): A HM approach is an important part of drug design that provides valuable insight into the 3D structure of protein targets using the template of a related protein. This method accelerates the process of structure-based drug discovery, providing a rational approach to understanding the interaction between drugs and their respective target proteins [5].

Molecular docking experiments involve computing how a small molecule (a potential drug) will bind to a target protein. In this way, researchers can evaluate binding affinity and determine the most favorable binding orientations for potential drug candidates [6].

Binding Energy Calculations of drug candidates can indicate whether they will have stronger and more favorable interactions [7]. By analyzing the free energy of binding, researchers can gain insights into the potential efficacy of a drug candidate and make informed decisions about its further development and validation [8].

Ligand-based drug design (LBDD) - the design of drugs by the knowledge of molecules, the following approaches can be used to find new drugs: Molecular dynamics, Quantitative structure-activity relationships (QSAR) modeling, Machine Learning, etc. [4, 9, 10].

Virtual Screening: Using computer models to screen large databases of chemical compounds could reduce the number of potential candidates for experimental testing, reducing the amount of time and resources spent on testing compounds that have the greatest chance of success [11, 12].

Optimization of Lead Compounds: To refine drug candidates before they are tested in the laboratory, researchers use computer modeling to optimize lead compounds' structures for better binding affinity, selectivity, and other desirable properties. It is an approach to improving lead compounds through a structure-activity relationship (SAR), where the structure of the lead compound is modified to identify key functional groups to its activity [13].

Drug Design: By using computer modeling, we can effectively screen potential drug candidates and predict their behavior in the body, ultimately leading to the development of safer and more effective treatments. Through this approach, pharmaceutical agents can be designed, optimized, and tailored by understanding the molecular target of a disease, in contrast to empirical approaches that work by trial and error [14, 15].

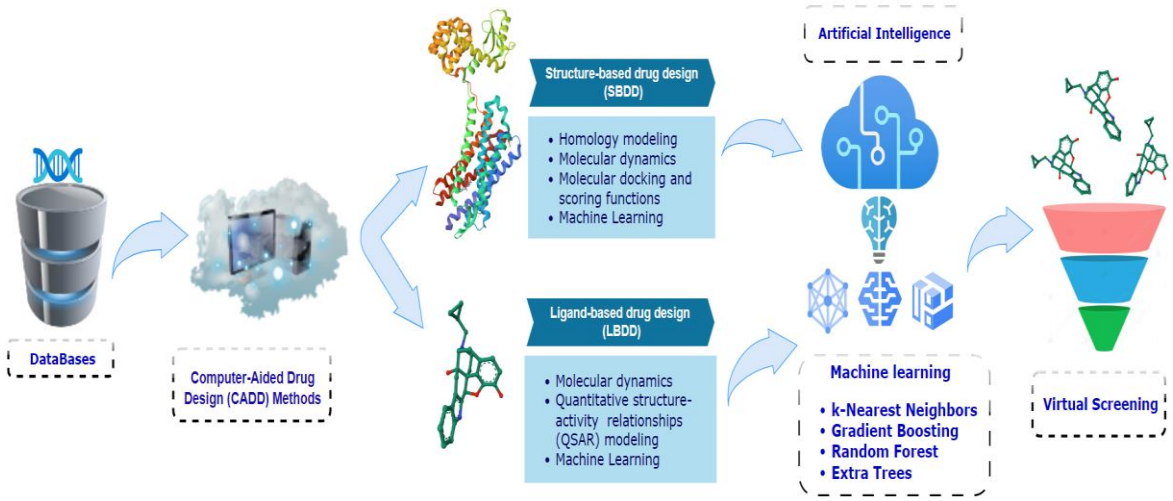

Fig.1. An illustration of how computer modeling, docking experiments, and machine learning contribute to the design of drugs

Machine and deep learning: Artificial Intelligence (AI) is also called machine intelligence in the field of computer science since machines are trained to perform tasks similar to those of humans [16]. The

concept of AI is related to disciplines like pattern recognition, image recognition [17], probability theory, statistics, machine learning [18], neural networks, etc. Computational intelligence provides a variety of approaches for analyzing, learning, and elucidating how AI plays a key role in the identification of numerous pharmaceuticals and automates the discovery of new drugs by an integrated approach [19, 20, 21, 22].

In the present research, we investigated tree models of Delta-opioid receptors (DOR) and delta-opioid ligands previously studied for their structure-activity relationships (SAR) in the following articles [23, 24, 25, 26]. We performed docking studies with the software GOLD and all optimization functions embedded in the toll [27, 28]. After docking studies, we calculated the binding for the compounds with the software MMV and used the optimization function MolDock [29]. In the study [26] we have searched for a relationship between docking results and the biological activity of the studied compounds. We found correlations between the values of biological activity of the ligands and the values of scoring function from the docking experiments. The discovered correlations underscore the significance of optimization function values and provide useful insights for the discovery and development of drugs. In publications [26, 30, 31] we applied the polynomial modeling approach to establish the relationship between docking results and the biological activity of delta-opioid ligands. We used polynomial function $z = f(x, y)$, where: z – values of the biological activity of the ligands; x– values of the scoring function from GOLD; y – values of the MolDock function from MMV.

This study aims to investigate the relationship between the structure and biological activity of delta-opioid ligands using the approaches of Machine Learning. Once again, we are looking for a relationship $z = f$) between the values of the biological activity of the ligands z, where the independent variables *x* are the docking results – the values of the different scoring functions and *y* are the values of the total energies – MolDock function. Our next step is to establish the nonlinear relationship using machine learning [32, 33].

2. Materials and methods

Our study used three models of the delta-opioid receptor (DOR) and delta-opioid ligands synthesized and analyzed in the following articles [23, 24, 25, 26]:
- a theoretical model of DOR (PDBe: 1ozc);
- a model of DOR with the crystal structure (PDBid: 4ej4) (https://www.rcsb.org/)
- a model of DOR, obtained by the Homology Modeling approach.

Table 1
Values of the data for delta-opioid ligands and three models of DOR

Ligands	**A theoretical model of DOR (PDBe: 1ozc)**			**Model of DOR, obtained by the Homology Modeling**			**Model of DOR (PDBid: 4ej4) with crystal structure**		
	Values of GoldScore function	Values of Total Energy	Biological effect	Values of ASP function	Values of Total Energy	Biological effect (IC_{50})	Values of Chem Score function	Values of Total Energy	Biological effect
[Cys(Bzl)^2-Leu^5]-enk	64,68	-107.022	9.3	20.26	-77.135	8.3	38.91	-170.657	9.3
[Cys(Bzl)^2-Met^5]-enk	81,49	-89.091	3.5	25.16	-98.91	9.53	35.19	-125.108	3.5
[Cys(O_2NH_2)^2-Leu^5]-enk	67,72	-97.619	29.2	22.66	-99.678	1.29	28.48	-118.805	29.2
[Cys(O_2NH_2)^2-Met^5]-enk	73,91	-91.246	7.3	26.18	-88.498	2.22	25.82	-87.343	7.3
[DCys(O_2NH_2)^2-Leu^5]-enk	74,73	-84.852	7.4	24.31	-66.115	11.4	31.84	-136.187	7.4
[DCys(O_2NH_2)^2-Met^5]-enk	75,13	-86.221	7.1	-12.82	897.265	75.96	31.55	-139.449	7.1

DPDPE	57,67	-109.709	30.2	19.58	-75.943	6.18	32.75	-100.702	30.2
[HCys(O_2NH_2)^2-Leu^5]-enk	68,43	-62.774	3.4	18.87	-90.567	31.92	26.55	-112.164	3.4
[HCys(O_2NH_2)^2-Met^5]-enk	78,65	-93.301	4.5	23.84	-80.137	16.09	29.23	896.877	4.5
[Leu^5]-enkephalin	73,42	-81.869	5.8	22.45	-104.149	11.45	31.62	-119.009	5.8
[Met^5]-enkephalin	73,26	-118.971	3.6	33.9	-112.752	18.91	32.22	-106.792	3.6

We performed docking studies with the software GOLD and all optimization functions embedded in the toll: GoldScore, ChemScore, ChemPLP, ASP Score [26, 34]. After that, the binding free energies of the studied delta-opioid ligands using the MMV program and the MolDock scoring function were computed. We used the Surface Fitting Tool of Matlab (http://www.mathworks.com) to model the data for three DOR models with polynomial function $z = f(x, y)$, where z represents the values of the biological activity of the ligands; x– values of the scoring function from GOLD; y – values of the MolDock function from MMV. The results obtained are presented in Table 1.

2.1. Machine learning techniques

In this subsection, we discuss the machine learning regression methods that we employ to obtain a functional relationship $z = f(x, y)$:

- **k-Nearest Neighbors (k-NN)** – it is a machine learning algorithm, based on a distance metric, like Euclidean or Manhattan distance; k-NN finds the k nearest neighbors to a given data point [35].
- **Gradient Boosting (GB)** – it is a powerful machine learning technique for constructing predictive models. It usually produces simple decision trees that make very few assumptions about the data [36, 37].
- **Random Forest (RF)** – it involves a collection of decision trees and is used for classification and regression tasks. The combination of individual tree predictions allows RF to provide accurate and reliable results, making it a popular choice across diverse applications in machine learning and predictive analytics [38].
- **Extremely Randomized Trees (Extra Trees, ET)** – it creates multiple trees, similar to RF algorithms, during the training phase across the entire dataset. With different sets of features, the ET constructs trees over each observation in the dataset during training [39].

3. Results and discussion

AI involves computer systems learning from input or past data. Scientists incorporate AI algorithms in the drug design and discovery process [7, 8]. This integration has significantly accelerated the identification of potential drug candidates, allowing researchers to analyze large amounts of data and predict molecular interactions more efficiently than ever before [9].

Computational modeling, based on AI and ML principles, is a way to identify and validate chemical compounds, target identification, peptide synthesis, and drug design. This approach allows for faster and more efficient screening of potential drug candidates, leading to the development of new and improved treatments for various diseases. Computational modeling can help in understanding the structure-activity relationship of chemical compounds, predicting their behavior in biological systems, and optimizing their pharmacokinetic properties [10, 11].

In previous publications [26, 31, 32], we presented polynomial modeling to find the structure–biological activity relationship for delta-opioid ligands and three models of DOR. We employed a polynomial function $z = f(x, y)$ in our studies, where: z – values of the biological activity of the ligands; x– values of the scoring function from GOLD; y – values of the MolDock function from MMV. Using the least squares method, we estimated the acquired polynomial model as a surface-fitting function. To assess suitability, statistical parameters are utilized, such as: SSE (Sum of Square Error), R^2 (Coefficient of

Determination), $AdjustedR^2$, $RMSE$ (Root Mean Square Error).

We found that the third-degree polynomial model has a good fitting and predictive ability to explain the complex relationship between the structure of compounds and their biological efficacy:

(1) DOR(ePDBid:1ozc): $R^2 = 1.0$, $SSE = 0.009207$, $adjustedR^2 = 0.9999$, $RMSE = 0.096$.
(2) DOR, obtained by homology modeling: $R^2 = 1.0$, $SSE = 0.2460$, $adjustedR^2 = 0.9999$
(3) DOR(PDBid:4ej4):$R^2 = 0.9990$,$SSE = 0.9631$,$adjustedR^2 = 0.9901$, $RMSE = 0.9814$

In the present study, we want to apply the artificial intelligence approach to model the structure–biological activity relationship of delta-opioid ligands and tree models of DOR. Our next step of this investigation is to establish nonlinear relationships using machine learning techniques. The data are collected and processed as previously described. Several machine learning algorithms are used and trained to accurately capture the nonlinear relationship: k-Nearest Neighbors, Gradient Boosting, Random Forest, and Extremely Randomized Trees (Extra Trees). This will enable us to make predictions and obtain valuable information from the data. Our research aims not only to uncover valuable insights into the Structure-Activity Relationship (SAR) but also to facilitate the exploration of novel compounds with enhanced therapeutic efficacy. The four machine learning algorithms are applied to construct a multivariate regression. The performance of each algorithm is evaluated using cross-validation, and the best performing algorithm is selected based on the mean squared error. After selecting the best algorithm, the multivariate regression model is trained on the dataset and used to make predictions on new data. The model's performance is then assessed using various metrics such as R-squared and mean absolute error to ensure its accuracy and reliability.

In the computation of indicators for each algorithm, we implemented a 10-fold cross-validation. Furthermore, a procedure for hyperparameter optimization through grid search was employed to identify the optimal values. The results obtained are presented in Table 2.

Table 2
Modeling by using machine learning techniques for three models of DOR

Regressor	**Best hyperparameters**	**SSE**	**R-square**	**RMSE**
A theoretical model of DOR (PDBe: 1ozc) [33]				
k-Nearest Neighbours	{'metric': 'minkowski', 'n_neighbors': 2, 'weights': 'uniform'}	351.0150	0.6397	5.6489
Gradient Boosting	{'max_depth': 2, 'n_estimators': 50}	7.9934	0.9918	0.8525
Random Forest	{'max_depth': 2, 'n_estimators': 50}	232.3306	0.7615	4.5958
Extra Trees	{'max_depth': 2, 'n_estimators': 100}	299.7497	0.6923	5.2202
Model of DOR, obtained by the Homology Modeling				
k-Nearest Neighbours	{'metric': 'minkowski', 'n_neighbors': 3, 'weights': 'uniform'}	2772.5207	0.3806	15.8760
Gradient Boosting	{'max_depth': 5, 'n_estimators': 50}	0.1858	0.999958	0.1300

Random Forest	{'max_depth': 2, 'n_estimators': 150}	689.4304	0.8460	7.9168
Extra Trees	{'max_depth': 2, 'n_estimators': 100}	586.8886	0.8689	7.3043
Model of DOR (PDBid: 4ej4) with crystal structure [34]				
k-Nearest Neighbours	{'metric': 'minkowski', 'n_neighbors': 5, 'weights': 'uniform'}	1223.4328	-0.2559	10.5461
Gradient Boosting	{'max_depth': 2, 'n_estimators': 50}	6.6080	0.9932	0.7751
Random Forest	{'max_depth': 4, 'n_estimators': 200}	228.5116	0.7654	4.5578
Extra Trees	{'max_depth': 2, 'n_estimators': 200}	763.3891	0.2163	8.3306

Cross-validation is a method employed to evaluate the performance of a machine-learning model on an external dataset, one that the model has not encountered during its training phase. In other words, it provides an assessment of how well the model can generalize its learnings to new and unfamiliar data. In our experiments, we employ the k-fold cross-validation technique, where the original sample is randomly divided into k subsamples of equal size. During this process, the model undergoes testing with one subsample, while the remaining k-1 subsamples are dedicated to training purposes. This entails testing the model with one subsample, with the other k-1 subsamples utilized for training. Throughout this procedure of cross-validation, each sub-sample is assigned as the validation data exactly once during each iteration. To generate a single estimate of model performance, the results of k-folds can be averaged or otherwise aggregated. To obtain a single estimate of model performance, the outcomes of k-folds can be averaged or aggregated in another way. This approach provides a more reliable and robust estimate of model performance compared to using a single train-test split. It also allows for a better understanding of the variability in the model's performance and can help identify potential issues such as overfitting or underfitting. Overall, using k-fold cross-validation can lead to more accurate and trustworthy assessments of a model's capabilities.

The goal of hyperparameter optimization is to find the best combination of hyperparameters and it is crucial for improving the performance of machine learning models. By systematically searching through different hyperparameter values, we can identify the combination that results in the highest accuracy, precision, or whatever metric is most important for the specific task. This process can involve techniques such as grid search, random search, genetic algorithms, etc. Hyperparameters in RF models may include how many decision trees are used in the forest, or what features are considered when a node is split, which are not learned from data. Hyperparameters are those that are not learned from the data and are set before the learning process begins. Unlike model parameters, hyperparameters are set before the learning process begins and are not updated during the learning process. Hyperparameters can significantly influence model performance, and thus it is crucial to choose them wisely.

We use the grid search strategy for this procedure, defining a set of possible values for each hyperparameter and checking the performance of every possible combination. This approach allows us to systematically explore the hyperparameter space and identify the combination that yields the best results.

Through a comprehensive exploration of all conceivable hyperparameter values, we guarantee that no potentially optimal configurations are overlooked. Although this approach may demand significant computational resources, it stands as a meticulous and dependable means of refining our models to achieve optimal performance (Table 2). The forecast surface for three models of DOR is presented in Figure 2.

As can be seen in the second column of Table 2, the most effective values of hyperparameters were

determined for each method. It is crucial for achieving the best performance and results. By carefully selecting the right hyperparameters, we can fine the model to improve its accuracy and generalization capabilities. In comparison to the polynomial modeling data performed in Matlab [30], they demonstrate comparable performance to that of a third-order polynomial, known as Poly33. The optimal method is Gradient Boosting. This approach is presented in Figure 3.

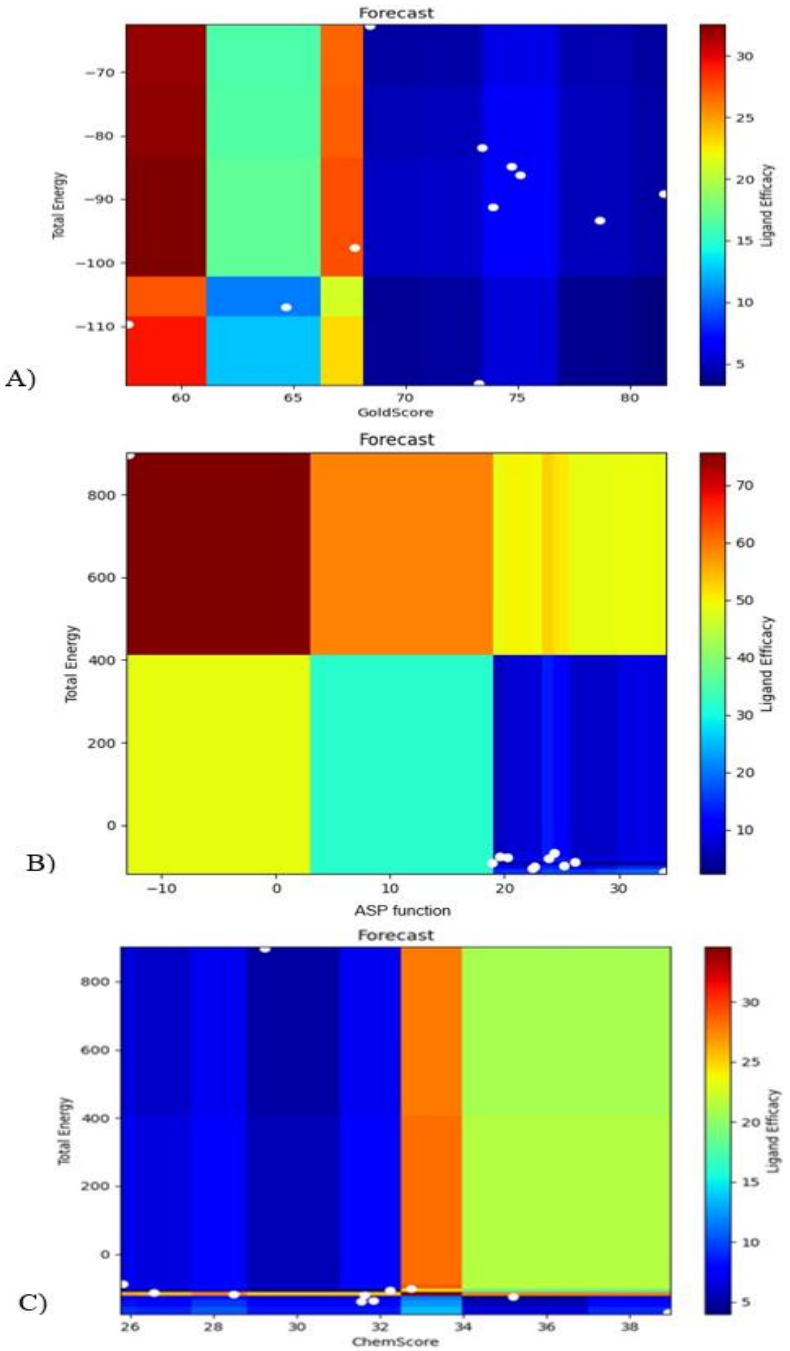

Fig. 2. The forecast surface for three models of DOR. A) a theoretical model of DOR (ePDBid:1ozc) and GoldScore scoring function, B) a model of DOR, obtained by homology modeling and ASP scoring function; C) a model of DOR (PDBid: 4ej4) with crystal structure and ChemScore scoring function

Fig. 3. Gradient Boosting algorithm

4. Conclusions

In this study, we explored the application of machine learning techniques to model the structure–biological activity relationship (SAR) of delta-opioid ligands across three delta-opioid receptor (DOR) models. Building upon previous work involving polynomial modeling, we incorporated advanced regression methods, including k-Nearest Neighbors (k-NN), Gradient Boosting (GB), Random Forest (RF), and Extremely Randomized Trees (Extra Trees), to establish nonlinear relationships between docking results, binding energies, and biological activity.

Our findings indicate that machine learning provides significant advantages over traditional modeling techniques by accurately capturing complex nonlinear relationships in SAR. Among the methods tested, Gradient Boosting demonstrated superior performance in terms of predictive accuracy and reliability, as confirmed by cross-validation and hyperparameter optimization. These results were consistent across all three DOR models, underscoring the robustness of the approach.

The integration of machine learning into the computational drug design process represents a transformative advancement [40]. By leveraging AI algorithms, we were able to identify patterns and correlations that would have been challenging to discern using conventional methods. This study highlights the potential of machine learning to enhance drug discovery by enabling faster, more precise predictions of biological activity, reducing the time and cost of experimental testing.

Future work will focus on extending this methodology to larger datasets and exploring additional machine learning algorithms to further optimize performance. Additionally, we plan to apply these models to other receptor-ligand systems to generalize their applicability and contribute to the development of novel therapeutic agents. The synergy between artificial intelligence and computational drug design has the potential to significantly optimize the drug discovery, paving the way for safer, more effective treatments for a wide range of diseases.

Acknowledgments

The authors would like to thank the anonymous reviewers, whose comments have improved the quality of the paper.

This research is supported by the Bulgarian National Science Fund under Project KP-06-M 62/1 "Numerical deterministic, stochastic, machine and deep learning methods with applications in computational, quantitative, algorithmic finance, biomathematics, ecology and algebra" from 2022.

References

[1] I. Muegge, A Bergner, J.M Kriegl, 2017. Computer-aided drug design at Boehringer Ingelheim, J. Comput. Aided Mol. Des. 31, 3, 275-285, https://doi.org/10.1007/s10822-016-9975-3

[2] M. Batool, B. Ahmad, S. Choi, 2019. A Structure-Based Drug Discovery Paradigm. International journal of molecular sciences, 20, 11, 2783. https://doi.org/10.3390/ijms20112783

[3] P. Aparoy, K. Kumar Reddy, P. Reddanna, 2012. Structure and ligand-based drug design strategies in the development of novel 5-LOX inhibitors. Curr. Med. Chem., 19, 22, 3763-3778, https://doi.org/10.2174/092986712801661112

[4] Tanos França, 2015. Homology modeling: an important tool for the drug discovery, Journal of Biomolecular Structure and Dynamics. 33, 8, 1780-1793, doi: 10.1080/07391102.2014.971429

[5] L. Ferreira, R. dos Santos, G. Oliva, A. Andricopulo, 2015. Molecular docking and structure-based drug design strategies. Molecules. 20:13384–13421. doi: 10.3390/molecules200713384.

[6] A. Schön, N. Madani, A. Smith, J. Lalonde, E. Freire, 2011. Some binding-related drug properties are dependent on thermodynamic signature. Chemical biology and drug design. 77, 3, 161–165. https://doi.org/10.1111/j.1747-0285.2010.01075.x

[7] Z. Cournia, B. Allen, W. Sherman, 2017. Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations, Journal of Chemical Information and Modeling 57, 12, 2911-2937. DOI: 10.1021/acs.jcim.7b00564

[8] V. Sharma, S. Wakode, H. Kumar, 2021. Structure-and ligand-based drug design: Concepts, approaches, and challenges. Chemoinformatics and bioinformatics in the pharmaceutical sciences, 27-53. https://doi.org/10.1016/B978-0-12-821748-1.00004-X

[9] C. Acharya, A. Coop, J. Polli, A. Mackerell Jr., 2011. Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. Curr. Comput. Aided Drug Des. 7, 1,10–22. https://doi.org/10.2174/157340911793743547

[10] F. Gentile, J. Yaacoub, J. Gleave, et al., 2022. Artificial intelligence–enabled virtual screening of ultra-large chemical libraries with deep docking. Nat Protoc 17, 672–697 https://doi.org/10.1038/s41596-021-00659-2

[11] T. Oliveira, M. Silva, E. Maia, A. Silva, A. Taranto, 2023. Virtual Screening Algorithms in Drug Discovery: A Review Focused on Machine and Deep Learning Methods. Drugs Drug Candidates. 2, 311-334. https://doi.org/10.3390/ddc2020017

[12] B. Neves, R. Braga, C. Melo-Filho, J. Moreira-Filho, E. Muratov, C. Andrade, 2018. QSAR-Based Virtual Screening:Advances and Applications in Drug Discovery.Front.Pharmacol. 9, 1275. https://doi.org/10.3389/fphar.2018.01275

[13] G. Sliwoski, S. Kothiwale, J. Meiler, E. Lowe, 2014. Computational methods in drug discovery. Pharmacological reviews. 66,1, 334-395. DOI: https://doi.org/10.1124/pr.112.007336

[14] K. Mak, M. Pichika, 2019. Artificial intelligence in drug development: Present status and future prospects. Drug Discov. Today. 24, 773–780. doi: 10.1016/j.drudis.2018.11.014.

[15] D. Poole, A. Mackworth, R. Goebel, 1998. Computational Intelligence: A Logical Approach. New York: Oxford University Press.

[16] C. Bishop, 2013. Model-based machine learning. Philos Trans. A Math. Phys. Eng. Sci. 371:20120222. doi: 10.1098/rsta.2012.0222.

[17] M. Vasileva, G. Prandzhev, D. Dimitrov, Z. Gorcheva, 2024. Current state of application of Artificial Intelligence in preoperative MRI assessment of endometrial cancer a mini review. Series on Biomechanics 38, 4,127-132. doi: 10.7546/SB.17.04.2024.

[18] A. Das, B. Lal, R. Manjunatha, 2022. Advances in Gravimetric Electronic Nose for Biomarkers Detection. Series on Biomechanics 36, 2,128-140. doi: 10.7546/SB.36.2022.02.13.

[19] W. Duch, K. Swaminathan, J. Meller, 2007. Artificial intelligence approaches for rational drug design and discovery. Curr. Pharm. Des. 13,1497–1508. doi: 10.2174/138161207780765954.

[20] K. Butler, D. Davies, H. Cartwright, O. Isayev, A. Walsh, 2018. Machine learning for molecular and materials science. Nature. 2018, 559, 547–555. doi: 10.1038/s41586-018-0337-2
[21] A. Jordan, 2018. Artifcial intelligence in drug design–the storm before the calm? ACS Med Chem Lett. https://doi.org/10.1021/acsmedchemlett.8b00500
[22] A. Goel, J. Davies, 2019 Artifcial intelligence. In: The Cambridge Handbook of Intelligence. Cambridge.
[23] N. Pencheva, A. Bocheva, E. Dimitrov, C. Ivancheva, R. Radomirov, 1996. [Cys(O2NH2)2]enkephalin analogues and dalargin: selectivity for delta-opioid receptors. Eur J Pharmacol. 23, 304(1-3):99-108. doi: 10.1016/0014-2999(96)00083-0.
[24] N. Pencheva, P. Milanov, L. Vezenkov, T. Pajpanova, E. Naydenova, 2004. Opioid profiles of Cys2-containing enkephalin analogues. Eur J Pharmacol. 13; 498(1-3):249-56. doi: 10.1016/j.ejphar.2004.07.059.
[25] P. Milanov, N. Pencheva, 2011. Theoretical hyperbolic model of a partial agonism: explicit formulas for affinity, efficacy and amplification. Serdica J. Computing, 5, 333-358.
[26] F. Sapundzhi, T. Dzimbova, P. Milanov, N. Pencheva, 2017. QSAR modelling and molecular docking studies of three models of delta opioid receptor J. Comput. Methods Molec. Design, 5, 2, 98-108.
[27] G. Jones, P. Willett, R.Glen, A. Leach, R. Taylor, 1997. Development and validation of a genetic algorithm for flexible docking, Journal of Molecular Biology, 267, 3, 727-748, https://doi.org/10.1006/jmbi.1996.0897.
[28] M. Verdonk, J. Cole, M. Hartshorn, C. Murray, R.Taylor, 2003. Improved protein-ligand docking using GOLD Proteins, 52, 609-623 doi: 10.1002/prot.10465.
[29] R. Thomsen, M. Christensen, 2006. A new technique for high-accuracy molecular docking. J Med Chem. 49, 11, 3315-3321. doi: 10.1021/jm051197e. PMID: 16722650.
[30] F. Sapundzhi, T. Dzimbova, 2019. A Study of QSAR based on Polynomial Modeling in Matlab. International Journal of Online and Biomedical Engineering 15, 15, 39–56. https://doi.org/10.3991/ijoe.v15i15.11566
[31] F. Sapundzhi, T. Dzimbova, P. Milanov, N. Pencheva, 2015. Determination of the relationship between the docking studies and the biological activity of delta-selective enkephalin analogues J. Comput. Methods Molec. Design, 5, 2, 98-108.
[32] F. Sapundzhi, M. Lazarova, T. Dzimbova, S.Georgiev, 2023. An application of some machine learning methods for biological data modeling. Journal of Physics: Conference Series. 2675, 1, 012021. doi:10.1088/1742-6596/2675/1/01202
[33] F. Sapundzhi, M. Lazarova, T. Dzimbova, S. Georgiev. A. Ivanova, 2023. A structure-activity relationship modelling of opioid compounds by using machine learning. Journal of Physics: Conference Series. 2675, 1, 012032. doi:10.1088/1742-6596/2675/1/012032
[34] F. Sapundzhi, K. Prodanova, M. Lazarova, 2019. Survey of the scoring functions for protein-ligand docking AIP Conference Proceedings.2172,100008,1-6 https://doi.org/10.1063/1.5133601
[35] O. Kramer, 2013. K-Nearest Neighbors. Dimensionality Reduction with Unsupervised Nearest Neighbors. Intelligent Systems Reference Library, 51. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-38652-7_2
[36] S. Piryonesi, T. El-Diraby, E. Tamer, 2020. Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index. Journal of Infrastructure Systems. 26, 1. 04019036. doi:10.1061/(ASCE)IS.1943-555X.0000512.
[37] T. Hastie, R. Tibshirani, J. Friedman, 2009. Boosting and Additive Trees. The Elements of Statistical Learning (2nd ed.). New York: Springer. 337–384.
[38] T. Austin, 2015. Exchangeable random measures, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, Ann. Inst. H. Poincaré Probab. Statist. 51, 3, 842-861.
[39] P. Geurts, D. Ernst, L. Wehenkel, 2006. Extremely randomized trees. Mach Learn 63, 3–42. https://doi.org/10.1007/s10994-006-6226-1.
[40] G. Kuchava, M. Mantskava, 2021. Brain blood flow control with artificial intelligence. Series on Biomechanics 35, 2, 73-78.

Information about the authors

Fatima Sapundzi – Ph. D., Associate Professor, Department of Communication and Computer Engineering, Faculty of Engineering, Neofit Rilski South-West University, 66 Ivan Myhailov Str., 2700 Blagoevgrad, BG; sapundzhi@swu.bg; ORCID: 0000-0002-5533-322X
Slavi Georgiev – Ph. D., Chief Assistant Professor, Department of Informational Modeling, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 8 Acad. Georgi Bonchev Str., 1113 Sofia, BG; sggeorgiev@math.bas.bg; ORCID: 0000-0001-9826-9603

Meglena Lazarova – Ph. D., Chief Assistant Professor, Faculty of Applied Mathematics and Informatics, Technical University of Sofia, 8 Kliment Ohridski Blvd., 1000 Sofia, BG; meglena.laz@tu-sofia.bg; ORCID: 0009-0001-8723-0014
Metodi Popstoilov – Ph. D., Associate Professor, Department of Communication and Computer Engineering, Faculty of Engineering, Neofit Rilski South-West University, 66 Ivan Myhailov Str., 2700 Blagoevgrad, BG; m_popstoilov@abv.bg; ORCID: