IJSRIS journal
INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH AND INNOVATIVE STUDIES
ISSN: 2820-7157
www.ijsrisjournal.com

November 2022 **Volume 1** **Number 3** **Pages 69-73**

Artificial Intelligence and Molecular Docking Focus on Selected Methods

Ibrahim Ahmed Hassan Alshuwaysh[1], Jawad Ali Abdullah Alhajji[1], Salwa Ali Abdullah Alrizq[2], Aqeel Mohammed Nasir Alsalman[3], Naif Saleh Jassim Alatiah[4], Ahmed Taher Ali Alaithan[5], Fatemah Mohammed Altolahiy[6], Amin Habib Saleh Albarrak[7], Ali Saleh Tahir Alsaleh[8].

1. Alhafayer PHC , Saudi Arabia.
2. Al-Shoqiq PHC, Saudi Arabia.
3. Hazm PHC, Saudi Arabia.
4. Al-Qurain PHC, Saudi Arabia.
5. Al-Mudayri Building , Saudi Arabia.
6. Prince Sultan Cardiac Center (PSCC) , Saudi Arabia.
7. Jalijla Health Care Center, Saudi Arabia.
8. Alrashdia PHC, Saudi Arabia.

ABSTRACT

Artificial intelligence (AI) has emerged as a powerful tool in drug discovery and development, revolutionizing the field of molecular docking. This review provides an overview of current approaches in molecular docking using AI techniques, including machine learning and deep learning algorithms. We discuss the application of AI in virtual screening, ligand-receptor binding prediction, and drug optimization, highlighting recent advancements, challenges, and future directions.

Keywords: Artificial intelligence, molecular docking, virtual screening, drug discovery, machine learning, deep learning.

Application of AI in Molecular Docking:

AI techniques, including machine learning and deep learning algorithms, have been widely applied in molecular docking to enhance the efficiency and accuracy of ligand-receptor binding prediction. Machine learning models learn from large datasets of molecular structures and binding affinities to predict the binding affinity between ligands and receptors(1). AI technologies can help analyze vast amounts of data, such as genomic, proteomic, and chemical information, to identify potential drug molecules and predict drug efficacy or toxicity(2). By analyzing complex datasets and identifying hidden patterns, machine learning (ML) or deep learning (DL) algorithms can find novel targets associated with multi-omics data and help search for novel chemical entities with biological activities. They have not only expedited the identification of potential drug candidates but have also proven invaluable in the process of drug repurposing (3). AI it predict potential new uses for existing drugs, a breakthrough that has the potential to accelerate the drug development process and reduce associated costs(3). This ability is particularly significant in addressing urgent medical needs, as repurposing existing drugs can bypass lengthy and costly phases of preclinical testing and safety evaluation. Moreover, AI has emerged as a key tool for personalized medicine by aiding the development of drugs that are tailored to individual patients' genetic profiles. In the future, the demand for AI in drug discovery is expected to grow as the technology becomes more advanced and many more data become available(4).

Machine learning in drug discovery by AI

Machine learning is a group of which is focus on the development of a logarithms that learn adapt and perform task through data processing and analysis(5). By identifying patterns, making predictions, and refining algorithms based on input data, ML allows machines to improve their prediction performance and decision-making capabilities autonomously over time. ML algorithms can be broadly categorized into three types: supervised learning, unsupervised learning, and reinforcement learning.

Support Vector Machine

Support Vector Machine (SVM) is a supervised machine learning algorithm that has been widely used in drug discovery and development for various applications, including virtual screening, quantitative structure-activity relationship (QSAR) modeling, and drug-target interaction prediction. Here's an overview of how SVM is used in drug discovery. SVM has

been employed in virtual screening to identify potential drug candidates from large compound libraries. In virtual screening, SVM models are trained on a dataset of known active and inactive compounds against a specific drug target. These models learn to discriminate between active and inactive compounds based on their molecular features, such as chemical structure, physicochemical properties, and molecular descriptors. Once trained, the SVM model can efficiently classify new compounds as potential drug candidates or non-candidates, thereby reducing the number of compounds that need to be experimentally tested(6). SVMs have established themselves as a significant tool in drug discovery due to their superior ability to analyze complex cheminformatics data. Their use extends to various tasks: they help in virtual screening processes(7),(8).

Naïve Bayes

Naïve Bayes is employed in virtual screening to prioritize compounds from large chemical libraries for experimental testing. In this application, Naïve Bayes classifiers are trained on a dataset of known active and inactive compounds against a specific drug target. The algorithm learns the probability distribution of features (e.g., molecular descriptors, fingerprints) associated with active and inactive compounds. When presented with new compounds, the Naïve Bayes classifier calculates the probability that each compound is active against the target based on its features. Compounds with high probabilities are considered potential drug candidates and are selected for further experimental validation(9). In drug design, the naïve Bayes algorithm has been widely applied, helping predict the biological activities of compounds, assisting in the early selection of promising candidates, and estimating results before laboratory experiments(10),(11). It can predict protein–protein(12) and drug–drug interactions(13), which is vital for understanding cellular pathways and managing polypharmacy, where patients take multiple drugs. This algorithm can also anticipate drug–target interactions, facilitating drug repurposing and side effect prediction(14)(15). In addition, it can classify compounds into specific categories quickly, although it operates on the assumption of feature independence, which may not always hold (16).

Random Forest

Random Forest is employed in virtual screening to prioritize compounds from large chemical libraries for experimental testing. In this application, Random Forest models are trained on a dataset of known active and inactive compounds against a specific drug target. The algorithm constructs an ensemble of decision trees, where each tree is trained on a subset of the training data with replacement. When presented with new compounds, the Random Forest model aggregates the predictions from individual trees to calculate the probability that each compound is active against the target. Compounds with high probabilities are considered potential drug candidates and are selected for further experimental validation (9).

Virtual Screening and Ligand Optimization

AI-driven molecular docking enables high-throughput virtual screening of large compound libraries to identify potential

drug candidates with high binding affinity and specificity for the target receptor. By rapidly screening millions of compounds, AI algorithms facilitate the identification of lead compounds for further experimental validation and optimization (17). Moreover, AI-based approaches can optimize the chemical structure of lead compounds to improve their binding affinity, pharmacokinetic properties, and safety profiles, thereby accelerating the drug discovery process (18). Despite significant advancements, AI-driven molecular docking still faces several challenges, including the need for large and diverse training datasets, the interpretability of AI models, and the generalization of predictions to novel ligand-receptor pairs. Future research directions include the development of hybrid AI models, integration of structural biology data, and incorporation of quantum mechanical calculations to improve the accuracy and reliability of AI-driven molecular docking [5].

Conclusion

In conclusion, AI-driven molecular docking holds tremendous promise for accelerating drug discovery and development. By leveraging machine learning and deep learning algorithms, AI enables rapid and accurate prediction of ligand-receptor binding, virtual screening of compound libraries, and optimization of lead compounds, thereby expediting the development of novel therapeutics for various diseases.

Authors contribution

All authors contribute equally in this work

Competing interest

None

References:

.1. Li H, Leung KS, Wong MH. Idock: A multithreaded virtual screening tool for flexible ligand docking. In: 2012 IEEE Symposium on Computational Intelligence and Computational Biology, CIBCB 2012. 2012.

2. Arabi AA. Artificial intelligence in drug design: algorithms, applications, challenges and ethics. Futur Drug Discov. 2021;3(2).

3. Gupta R, Srivastava D, Sahu M, Tiwari S, Ambasta RK, Kumar P. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. Mol Divers. 2021;25(3).

4. Chen W, Liu X, Zhang S, Chen S. Artificial intelligence for drug discovery: Resources, methods, and applications. Vol. 31, Molecular Therapy Nucleic Acids. 2023.

5. Cerchia C, Lavecchia A. New avenues in artificial-intelligence-assisted drug discovery. Vol. 28, Drug Discovery Today. 2023.

6. Ballester PJ, Mitchell JBO. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. Bioinformatics. 2010;26(9).

7. Warszycki D, Struski Ł, Śmieja M, Kafel R, Kurczab R. Pharmacoprint: A Combination of a Pharmacophore Fingerprint and Artificial Intelligence as a Tool for Computer-Aided Drug Design. J Chem Inf Model. 2021;61(10).

8. Jayaraj PB, Jain S. Ligand based virtual screening using SVM on GPU. Comput Biol Chem. 2019;83.
9. Gabel J, Desaphy J, Rognan D. Beware of machine learning-based scoring functions-on the danger of developing black boxes. J Chem Inf Model. 2014;54(10).
10. Hu Y, Lu Y, Wang S, Zhang M, Qu X, Niu B. Application of Machine Learning Approaches for the Design and Study of Anticancer Drugs. Curr Drug Targets. 2018;20(5).
11. Kaiser TM, Burger PB, Butch CJ, Pelly SC, Liotta DC. A Machine Learning Approach for Predicting HIV Reverse Transcriptase Mutation Susceptibility of Biologically Active Compounds. J Chem Inf Model. 2018;58(8).
12. Hu J, Zhou L, Li B, Zhang X, Chen N. Improve hot region prediction by analyzing different machine learning algorithms. BMC Bioinformatics. 2021;22.
13. Celebi R, Uyar H, Yasar E, Gumus O, Dikenelli O, Dumontier M. Evaluation of knowledge graph embedding approaches for drug-drug interaction prediction in realistic settings. BMC Bioinformatics. 2019;20(1).
14. Azlim Khan AK, Ahamed Hassain Malim NH. Comparative Studies on Resampling Techniques in Machine Learning and Deep Learning Models for Drug-Target Interaction Prediction. Vol. 28, Molecules. 2023.
15. Saha S, Chatterjee P, Halder AK, Nasipuri M, Basu S, Plewczynski D. ML-DTD: Machine Learning-Based Drug Target Discovery for the Potential Treatment of COVID-19. Vaccines. 2022;10(10).
16. Hou T, Bian Y, McGuire T, Xie XQ. Integrated multi-class classification and prediction of gpcr allosteric modulators by machine learning intelligence. Biomolecules. 2021;11(6).
17. Jiménez J, Škalič M, Martínez-Rosell G, De Fabritiis G. KDEEP: Protein-Ligand Absolute Binding Affinity Prediction via 3D-Convolutional Neural Networks. J Chem Inf Model. 2018;58(2).
18. Kumari M, Subbarao N. Convolutional neural network-based quantitative structure-activity relationship and fingerprint analysis against inhibitors of anthrax lethal factor. Future Med Chem. 2023;15(10).