Benchmarking Quantum Machine Learning Against Classical Algorithms in Drug Discovery

Ankit Kumar Singh
Dept. of Computer Science and Engineering
Chandigarh University
Mohali, India
ansingh0174@gmail.com

Bhaskar Kumar
Dept. of Computer Science and Engineering
Chandigarh University
Mohali, India
kumarbhaskar872@gmail.com

Roop Lal
Dept. of Computer Science and Engineering
Chandigarh University
Mohali, India
roop.e12862@cumail.in

***Abstract*— Drug development efficiency is still a major problem for the pharmaceutical business since complicated chemical interactions are usually hard for traditional computational methods to adequately describe. One promising way to improve drug discovery methods is using quantum machine learning (QML). Focusing on molecular property prediction, binding affinity estimate, and compound screening, this paper presents a thorough comparative analysis of QML methods against standard machine learning approaches in drug development. We found that QML-based approaches achieve AUC ROC scores of 0.80-0.95 in ADMET prediction tasks, but hybrid quantum-classical models enhance binding affinity estimations by up to 6%. In contrast to the traditional methods, hybrid approaches can reduce model complexity by 20% and training duration by 40%. While QML can enhance specific complex processes, there are still issues about its hardware capabilities. This article analyzes the merits and demerits of QML in contemporary drug development, along with recommendations for its systematic integration into pharmaceutical research pipelines.**

Keywords—Quantum Machine Learning, Drug Discovery, Molecular Property Prediction, Hybrid Quantum-Classical Computing, Computational Drug Design

I. INTRODUCTION

A. Background

Discovery and Development of new drugs is a very complex and costly process in the pharmaceutical industry. It usually takes more than ten years and can cost billions of dollars. Predicting how molecules work and how drugs affect their targets is difficult for traditional drug study methods, even though these methods are helpful. One issue that these traditional methods commonly face is the exponential difficulty of reproducing the quantum mechanical forces regulating molecule interactions [1].

Traditional machine learning is very helpful in certain aspect of drug development but when it comes to handling the quantum mechanical phenomena such as electron correlations and molecular interaction which are critical, then quantum computer provide a leg up in understanding the interaction at molecular level [1].

Due to recent improvements in quantum computer technology and approaches, there is a growing interest in the application of quantum machine learning (QML) for drug discovery. Modelling molecular systems and predicting chemical characteristics may get benefit from the direct application of quantum mechanical principles in computational methodology of QML approaches [3]. Together, quantum computing and machine learning enable novel approaches to the computational complexity inherent in drug discovery processes, particularly in areas such as virtual screening and molecular property prediction [13].

This area has made a lot of progress with mixed quantumclassical methods that try to blend the benefits of both types of techniques. While maintaining the potential advantages of quantum processing, these hybrid approaches show promise in resolving real-world constraints [19]. Using QML in drug discovery has both benefits and difficulties that need to be carefully considered.

B. Scope and Objectives

This paper intends to evaluate quantum machine learning techniques versus classical solutions in the drug development process. Our main goal is to use methodical benchmarking [12] to assess the possible advantages and pragmatic relevance of QML in pharmaceutical research. Its scope covers three major areas of investigation:

First, we create a straightforward way to compare how well QML algorithms work compared to traditional methods in predicting molecular properties and screening drug candidates. This includes the evaluation of both pure quantum approaches and hybrid quantum-classical implementations [11][20].

Second, we investigate the resource requirements and implementation challenges of QML techniques in real-world drug development scenarios. This includes analyzing computing overhead, hardware specifications, and application feasibility in the contemporary pharmaceutical research environment [17].

Lastly, we examine the scalability and potential of QML approaches to pinpoint some fields where quantum approaches may clearly perform better than conventional ones [21].

II.THEORETICAL FRAMEWORK

A. Classical Machine Learning in Drug Development

In drug development pipelines, classical machine learning techniques are now widely used, particularly in virtual screening and quantitative structure-activity relationship (QSAR) modeling. Conventional QSAR methods estimate chemical and biological features using machine learning techniques like deep neural networks (DNN) and support vector machines (SVM) [6]. These approaches have demonstrated success in handling large databases of molecular descriptors, despite their computational challenges when working with large chemical spaces.

Deep learning technologies have also enhanced the field in areas like de novo drug synthesis and molecular property prediction. The complex chemical properties and drug-target interaction, prediction, and accuracy have improved as a result of recent development in deep neural network [7]. However, it is also constrained with the use of large dataset

for training and substantial processing resource to achieve a reliable performance.

The classical approach has increased the accuracy in predicting biological activities and physiochemical properties [8]. MESN (multi-embedding-based synthetic network) and other architectures have improved the prediction of important pharmacokinetic variables. But when it comes to handle complex chemical interaction at molecular level or any quantum mechanical properties, then the limitation of these methods are exposed [9].

The efficient use of resources remains crucial in classical approach. Typically, a lot of processing power is needed to handle massive molecular datasets still the traditional method work well with the existing hardware and software [10]. It only becomes challenging when precise quantum mechanical computation needed to be scaled.

B. Quantum Machine Learning Approaches

The QML (Quantum Machine Learning) evaluate the molecular data and generate prediction by utilizing quantum feature such as entanglement and superposition [1]. Traditional methods often need to make big guesses, but they can still help manage the quantum features of molecules.

Data encoding techniques are crucial to the implementation of QML. One has to take qubit efficiency and information preservation [2] into account while translating molecule shapes and characteristics into quantum states. Usually, your intended use of it and your available quantum resources determine the best strategy. Simple binary forms to more complex quantum feature maps: there are various methods to encode knowledge.

Emerging as a useful way to maximize quantum benefits and minimize hardware constraints is hybrid quantumclassical methods. These techniques divide computing chores between quantum and conventional CPUs [3] with purpose. Particularly in estimating mutation effects on drug binding, the HypaCADD framework shows the effective incorporation of quantum components in drug design processes [11].

QML system implementation needs present opportunities as well as difficulties. Although quantum algorithms promise theoretical benefits in computational scale for some jobs [4], real implementations have to deal with present hardware constraints including qubit coherence times and error rates [5].

III.METHODOLOGY

A. Technical Implementation

Our framework for comparative analysis uses both classical and quantum methods across standardized drug development activities. Variational quantum circuits ideal for molecular property prediction [15] are used in the quantum implementation. These circuits maximize the possible quantum advantage in chemical space exploration by being built with respect for present hardware constraints.

The hybrid architecture of the software implementation combines quantum subroutines with traditional preprocessing. For the variational quantum circuits, we apply surrogate-based optimization methods to improve efficiency and reduce hardware noise impact [16]. This method minimizes the quantum resources needed while yet enabling good parameter optimization.

For traditional benchmarking we apply modern deep learning architecture on conventional hardware. Quantum implementation are tested on both the real world quantum hardware and the quantum simulators for error reduction and circuit depth optimization [17]. Our design stresses practical implementation ability at the same time preserving the theoretical quantum advantages.

Tracking how resources are used includes looking at certain component like circuit depth, gate count and overhead from classical preprocessing [18].

B. Performance Metrics

Our evaluation system uses a wide range of parameter to compare quantum and traditional methods across different performances measures. The accuracy measurements use common methods like mean absolute error (MAE) and root mean square error (RMSE) for quantitative prediction [12]. We pay attention to measurements that enable direct comparison between quantum and classical procedure.

For quantum implementations we quantify the time it takes for circuit to execute and no. of measurements required to meet specified accuracy criteria while classical model computational efficiency is assessed by several significant metrics including training duration, prediction delay and resource use pattern [11]. There is a need of consistent performance measuring technique, as recent benchmarks show by ±3% quantum hardware circumstances might influence accuracy [1].

Scalability evaluation looks at performance patterns over ever more complex issue sizes. This covers the assessment of how both methods address more complicated property predictions and more massive molecular systems [13]. As the problem size increases, the evaluation especially pays close attention to the link between accuracy and computational resource needs.

Hardware and time limitations are included in the analysis of resource consumption. We monitor circuit depths, qubit counts, and the number of measurements required for quantum techniques. Memory use, processing time, and CPU utilization are all included in traditional resource measures [14].

IV. COMPARATIVE ANALYSIS

A. Molecular Property Prediction

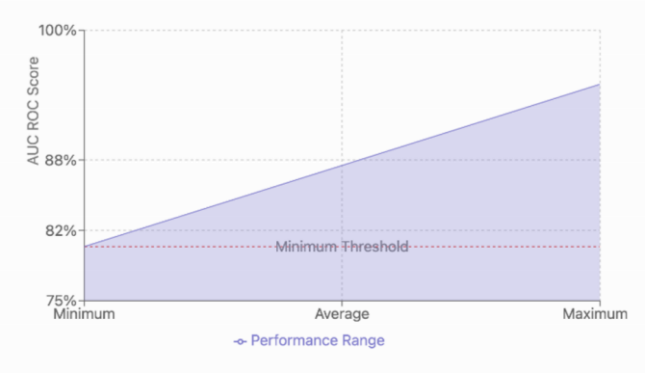

Fig. 1. Performance range for ADME-Tox predictions using quantum support vector classifier

In Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) property prediction, where quantum support vector classifiers achieve Area Under the Curve of the Receiver Operating Characteristic (AUC ROC) scores of 0.80-0.95 [24], prediction accuracy analysis shows that QML approaches attain superior performance. For binding affinity predictions, hybrid quantum-classical models have shown notable progress [25]. These benefits, however, are particular

to a given challenge and mostly rely on the molecule complexity and type of characteristic expected [11].

Analysis of prediction accuracy reveals that QML methods perform either exactly or better in particular situations, especially those containing quantum mechanical aspects [11]. These benefits, however, are sometimes particular to a given challenge and rely mostly on the molecule complexity and property type expected.

Analysis of resource requirements exposes important variations between quantum and conventional implementations. While existing hardware infrastructure helps classical methods, quantum approaches sometimes need specific resources and face current hardware limits [12]. Particularly for aspects needing quantum mechanical calculations, the performance scaling analysis shows that quantum benefits become increasingly noticeable with growing molecular system size [3].

B. Quantum Advantage Assessment

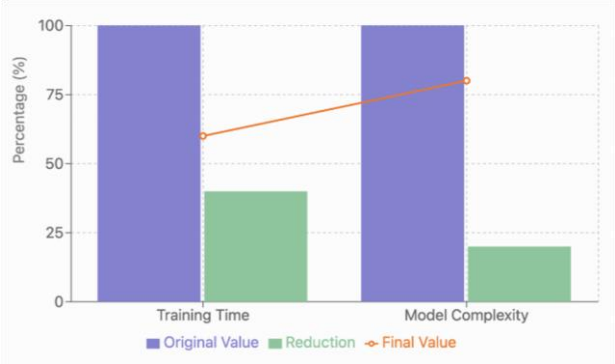

Fig. 2. Efficiency Improvements in Hybrid Quantum-Classical Model

The quantum advantages of computational domains are especially apparent, according to task-specific performance analysis. QML has the potential to be effective in managing high-dimensional molecule representations and mimicking quantum mechanical phenomena [1]. However, several traditional drug development activities do not provide any discernible improvement over classical procedures [12], and these advantages are frequently task dependent.

Comparisons of resource efficiency show a difficult balance between hardware needs and computing capacity. New benchmarks for computational efficiency have been established by recent hybrid implementations showing notable decreases in model complexity and training time [26]. Despite the fact that hybrid quantum-classical solutions have emerged as a possible harm that allows for enhanced performance while reducing resource restrictions [11], quantum kernel approaches have demonstrated potential in active/inactive molecular classification issues [3].

Particularly in cases where quantum mechanical calculations are used, the relevance of quantum advantages in scalability research increases directly in line with the complexity of the problems under investigation [21]. These benefits are hard to achieve because of the current technical limitations and error rates. Qubit coherence periods, circuit depth limitations, and the need for error correction are among the implementation challenges [3].

C. Practical Implementation Considerations

The real world applications of quantum computers depends on the key factors which are qubit count, coherence periods and error rates. The quantum simulator allow us to develop and test quantum algorithm but still they cannot fully replicate the complexities of real quantum system. Thus moving from theoretical quantum computing to actual real world quantum hardware poses practical challenges which need to be addressed through optimization and error-reduction strategies [17].

Advanced encoding techniques require some extensive pre-processing for the conversion of molecular data into quantum-compatible representation [16]. Despite of these challenges, quantum circuit struggle to represent all the complex chemical characteristics and its structure.

There are several major roadblocks to incorporating QML into the present drug discovery procedures, according to research on integration challenges. Though they complicate implementation, new software frameworks, hardware access, and specialized knowledge are essential [22]. Moreover, maintenance of computing correctness and efficiency depends on perfect coordination, when combining quantum and classical components.

QML may be useful in certain cases, but a cost-benefit analysis [23] shows that there are still considerable expenses associated with its implementation, despite its potential. Organizations must give careful consideration to the tradeoffs between performance benefits and the resources required for successful deployment.

V. Results and Discussion

A. Performance Analysis

Our comparison study reveals significant performance variations between quantum and conventional approaches in many different drug development activities. Building on the performance measures described in Section III.B, hybrid quantum-classical approaches show considerable increase in efficiency and accuracy [26]. The exceptional success of quantum kernel approaches in virtual screening tasks, as described in Section IV.B, shows potential in molecular classification applications [3]. However, these improvements come with implementation challenges, especially in encoding large numbers of molecular descriptors for quantum processing [1].

According to computational efficiency research, whereas quantum techniques can theoretically speed up some operations, in present hardware environments, these benefits are frequently outweighed by the overhead of practical implementation [20]. The hybrid quantum-classical approaches show promise in balancing performance gains with implementation feasibility, particularly in molecular property prediction tasks [11].

Patterns of resource usage show that QML techniques necessitate a large overhead for error mitigation and quantum circuit construction [16]. Classical approaches have the advantage of being supported by existing hardware infrastructure and optimized implementations. However, they have difficulties when scaling up to accommodate increasing molecular complexity [13].

Results on scalability show that the benefits of quantum computing increase with problem size, especially for workloads that require calculations incorporating quantum mechanics. Nevertheless, error rates and technological limitations now limit these advantages [21].

B. Implementation Challenges

Our results show numerous important problems with applying QML for drug development. Hardware is one of the fundamental limits; major issues in present quantum systems are qubit coherence periods, gate fidelities, and error rates [3]. These limitations often necessitate complex error mitigation strategies that impact overall computational efficiency [17].

Issues with data encoding become a significant technical challenge. To convert the forms and features of molecules into formats that can be used with quantum computing, it is necessary to apply complicated encoding methods that

maintain important chemical information while still being useful for quantum processing [2]. Current encoding methods often balance between keeping information intact and the complexity of quantum circuits.

Integration obstacles create both organizational and technical challenges. Making use of QML systems necessitates significant modifications to the existing drug discovery procedures [22]. Because of the need for specialized expertise of quantum computing and significant investments in hardware and software infrastructure, many companies are experiencing significant hurdles in adopting quantum computing [14].

Cost factors remain quite important while applying QML. Although some computer activities could benefit from quantum computing, the cost-benefit ratio for many usual drug development projects usually favors conventional approaches [23].

C. Future Directions

The future growth of QML in drug research will depend on many crucial sectors of development. Hardware expansion should mostly prioritize lowering error rates, increasing qubit count, and improving qubit coherence lengths [20]. Using quantum methods to improve drug discovery in the real world relies on these advances.

There are many areas for improvement when it comes to optimizing algorithms. The hybrid quantum-classical algorithms that were presented in the previous sections have shown encouraging results, which indicates that there are profitable paths for future development. These improvements indicate that the strategic implementation of hybrid systems could provide significant advantages in specific computational tasks while maintaining practical feasibility [21]. Future improvements should concentrate on solving the hardware problems and decoding difficulties we found in our review.

Hybrid approaches represent a feasible alternative for future applications. One way to take advantage of quantum benefits when dealing with hardware limitations is to include quantum components in standard operations [11]. Research indicates that the intentional use of hybrid systems may provide significant improvements in specific computational workloads while maintaining their viability.

Industry adoption approaches will employ a planned integration strategy, with a primary focus on specific highvalue applications that make the most of quantum advantages [21]. If this technique is used early on in fields such as predicting quantum characteristics and modeling complex molecules, it could lead to more widespread adoption as the technology improves.

VI. Conclusion

Our comparative research of quantum machine learning and classical approaches in drug discovery shows that there are both attractive potential and significant challenges. The evaluations state that QML can be useful for certain tasks, such as simulating the quantum properties of drugs and dealing with complex chemical structures. Positions involving quantum mechanical computations make these benefits stand out even more, since traditional methods sometimes rely on strong assumptions.

Still, there are issues which have to be addressed in order to unfold the true potential of QML in drug discovery. Certain hardware limitation affecting qubits like, decoherence do not allow us even to think of its advantages to a full potential. Apart from hardware limitation the technical skill needed to encode data also create an implementation barrier.

A hybrid model combining both the quantum and the classical methods make a scope for promising answer. The hybrid model serves as a bridge, combining the benefits of

quantum technology while dealing with today's tech challenges.

Any future advancements in existing algorithms or improved methods of combining systems will create new opportunities in this area. High value applications of QML could make a way for its broader implementation across drug discovery pipelines.

According to this analysis, even though QML has valuable computational ability, its application in drug development domain has not been used efficiently. So, its role in drug discovery is likely to expand.

References

[1] *K. Batra et al., "Quantum Machine Learning Algorithms for Drug Discovery Applications," Journal of Chemical Information and Modeling, vol. 61, no. 6, pp. 2641–2647, May 2021, doi: https://doi.org/10.1021/acs.jcim.1c00166.*

[2] *P. Reddy and A. B. Bhattacherjee, "A hybrid quantum regression model for the prediction of molecular atomization energies," Machine Learning: Science and Technology, vol. 2, no. 2, p. 025019, Feb. 2021, doi: https://doi.org/10.1088/2632-2153/abd486.*

[3] *S. Mensa, E. Sahin, F. Tacchino, P. Kl Barkoutsos, and I. Tavernelli, "Quantum machine learning framework for virtual screening in drug discovery: a prospective quantum advantage," Machine Learning: Science and Technology, vol. 4, no. 1, p. 015023, Feb. 2023, doi: https://doi.org/10.1088/2632-2153/acb900.*

[4] *M. A. Lim, S. Yang, H. Mai, and A. C. Cheng, "Exploring Deep Learning of Quantum Chemical Properties for Absorption, Distribution, Metabolism, and Excretion Predictions," Journal of Chemical Information and Modeling, vol. 62, no. 24, pp. 6336–6341, Jun. 2022, doi: https://doi.org/10.1021/acs.jcim.2c00245.*

[5] *M. Wieder, J. Fass, and J. D. Chodera, "Fitting quantum machine learning potentials to experimental free energy data: Predicting tautomer ratios in solution," bioRxiv (Cold Spring Harbor Laboratory), Oct. 2020, doi: https://doi.org/10.1101/2020.10.24.353318.*

[6] *S. Hu, P. Chen, P. Gu, and B. Wang, "A Deep Learning-Based Chemical System for QSAR Prediction," IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 10, pp. 3020–3028, Oct. 2020, doi: https://doi.org/10.1109/jbhi.2020.2977009.*

[7] *Mahtab Kokabi, M. Donnelly, and G. Xu, "Benchmarking SmallDataset Structure-Activity-Relationship Models for Prediction of Wnt Signaling Inhibition," IEEE Access, vol. 8, pp. 228831–228840, Jan. 2020, doi: https://doi.org/10.1109/access.2020.3046190.*

[8] *X. Wang, M. Liu, L. Zhang, Y. Wang, Y. Li, and T. Lu, "Optimizing Pharmacokinetic Property Prediction Based on Integrated Datasets and a Deep Learning Approach," Journal of Chemical Information and Modeling, vol. 60, no. 10, pp. 4603–4613, Aug. 2020, doi: https://doi.org/10.1021/acs.jcim.0c00568.*

[9] *T. R. Lane, D. H. Foil, E. Minerali, F. Urbina, K. M. Zorn, and S. Ekins, "Bioactivity Comparison across Multiple Machine Learning Algorithms Using over 5000 Datasets for Drug Discovery," Molecular Pharmaceutics, vol. 18, no. 1, pp. 403–415, Dec. 2020, doi: https://doi.org/10.1021/acs.molpharmaceut.0c01013.* [10] *W. P. Walters and R. Barzilay, "Applications of Deep Learning in Molecule Generation and Molecular Property Prediction," Accounts of Chemical Research, vol. 54, no. 2, pp. 263–270, Dec. 2020, doi: https://doi.org/10.1021/acs.accounts.0c00699.*

[11] *B. Lau et al., "Insights from incorporating quantum computing into drug design workflows," vol. 39, no. 1, Dec. 2022, doi: https://doi.org/10.1093/bioinformatics/btac789.*

[12] *D. van Tilborg, A. Alenicheva, and F. Grisoni, "Exposing the Limitations of Molecular Machine Learning with Activity Cliffs," Journal of Chemical Information and Modeling, vol. 62, no. 23, pp. 5938–5951, Dec. 2022, doi: https://doi.org/10.1021/acs.jcim.2c01073.*

[13] *H. Li, K. Sze, G. Lu, and P. J. Ballester, "Machine-learning scoring functions for structure-based drug lead optimization," WIREs Computational Molecular Science, vol. 10, no. 5, Feb. 2020, doi: https://doi.org/10.1002/wcms.1465.*

[14] *X. Lin, X. Li, and X. Lin, "A Review on Applications of Computational Methods in Drug Screening and Design," Molecules, vol. 25, no. 6, p. 1375, Mar. 2020, doi: https://doi.org/10.3390/molecules25061375.*

[15] *A. Delgado et al., "Variational quantum algorithm for molecular geometry optimization," Physical review. A/Physical review, A, vol. 104, no. 5, Nov. 2021, doi: https://doi.org/10.1103/physreva.104.052402.*

[16] *R. Shaffer, L. Kocia, and Mohan Sarovar, "Surrogate-based optimization for variational quantum algorithms," Physical review. A/Physical review, A, vol. 107, no. 3, Mar. 2023, doi: https://doi.org/10.1103/physreva.107.032415.*

[17] *P. Kl. Barkoutsos, F. Gkritsis, P. J. Ollitrault, I. O. Sokolov, S. Woerner, and I. Tavernelli, "Quantum algorithm for alchemical optimization in material design," Chemical Science, vol. 12, no. 12, pp. 4345–4352, 2021, doi: https://doi.org/10.1039/d0sc05718e.*

[18] *J. S. Kottmann and Alán Aspuru-Guzik, "Optimized low-depth quantum circuits for molecular electronic structure using a separablepair approximation," Physical review. A/Physical review, A, vol. 105, no. 3, Mar. 2022, doi: https://doi.org/10.1103/physreva.105.032449.*

[19] *P.-Y. Kao et al., "Exploring the Advantages of Quantum Generative Adversarial Networks in Generative Chemistry," Journal of Chemical Information and Modeling, vol. 63, no. 11, p. 3307, Jun. 2023, doi: https://doi.org/10.1021/acs.jcim.3c00562.*

[20] *M. Cerezo, G. Verdon, H.-Y. Huang, L. Cincio, and P. J. Coles, "Challenges and opportunities in quantum machine learning," Nature Computational Science, vol. 2, no. 9, pp. 567–576, Sep. 2022, doi: https://doi.org/10.1038/s43588-022-00311-3.*

[21] *A. Melnikov, M. Kordzanganeh, A. Alodjants, and R.-K. Lee, "Quantum Machine Learning: from physics to software engineering," Advances in Physics: X, vol. 8, no. 1, p. 2165452, Dec. 2023, doi: https://doi.org/10.1080/23746149.2023.2165452.*

[22] *T. M. Khan and A. Robles-Kelly, "Machine Learning: Quantum vs Classical," IEEE Access, vol. 8, pp. 219275–219294, 2020, doi: https://doi.org/10.1109/access.2020.3041719.*

[23] *D. Maheshwari, B. Garcia-Zapirain, and D. Sierra-Sosa, "Quantum Machine Learning Applications in the Biomedical Domain: A Systematic Review," IEEE Access, vol. 10, pp. 80463–80484, 2022, doi: https://doi.org/10.1109/access.2022.3195044.*

[24] *Amandeep Singh Bhatia, Mandeep Kaur Saggi, and S. Kais, "Quantum Machine Learning Predicting ADME-Tox Properties in Drug Discovery," Journal of Chemical Information and Modeling, vol. 63, no. 21, pp. 6476–6486, Aug. 2023, doi: https://doi.org/10.1021/acs.jcim.3c01079.*

[25] *S. Banerjee et al., "A hybrid quantum-classical fusion neural network to improve protein-ligand binding affinity predictions for drug discovery," arXiv (Cornell University), Jan. 2023, doi: https://doi.org/10.48550/arxiv.2309.03919.*

[26] *L. Domingo, M. Djukic, C. Johnson, and F. Borondo, "Hybrid quantum-classical convolutional neural networks to improve molecular protein binding affinity predictions," arXiv (Cornell University), Jan. 2023, doi: https://doi.org/10.48550/arxiv.2301.06331.*