ICEST-2K24 [International Conference on Engineering, Science and Technology]
In Association with
International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN - 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)

Pharmaceutical Data Optimization Using Quantum Machine Learning

Prof. Jalindar N. Ekatpure[1], Miss. Rupali B. Gavali[2], Mr. Swastik Y. Padasalkar[2], Miss. Prajakta I. Kale[2], Miss. Pallavi D. Jadhav[2]

[1]Assistant Professor, [2]Student

Department of Computer Engineering, S. B. Patil College of Engineering, Indapur, Maharashtra, India

ABSTRACT

In this few years new diseases like COVID-19 and its variants are emerging. There needs deeper aspects involved in improving the drugs .Generic medicine companies like Pfizer, hindustanUniliver had done many clinical trials on patients who were acyually infected by COVID-19 .It is also important knowing the effects of drug on a mass population before making it into production. So emerges the need for optimization in drugs. The previous information or datasets from the pharma companies, hospitals and clinics can be studies and we can also predict how much safe the orally consumed drugs are safe for the patient . Doing so will definitely ease the control over the effect of drug whether it is harmful or not depending upon its severity or risk factor. This encourages the development of a standardised metric that is well accepted by many i.e. nothing but the WHO (World Health Organization) accepted scores which will eventually help in categorising the drug compositions. Everytime the new information is updated into the live dataset, the 'Risk Assesment' team can analyse the risk, bioactivity on the group of drugs. This can also adher the business intelligence of the overall system. Here comes the quantum advantage, a revolutioning promise of quantum algorithms in solving the optimization or combinatorial problems in managing complex relationships among the data. That tools will help to reduce the time consumption. This paper introduces new approach for "pharmaceutical data optimization using quantum machine learning". Quantum algorithms are thoroughly designed for performing the machine learning operations harnessing the power of Quantum Computing on a NISQ based quantum computing.This is based on qubits rather than classical bits in comparison with classical bits.

Index Terms—Pharmacy, Quantum Machine Learning, Optimization ,RiskAssessment,Quantum algorithms, NISQ (Noisy Intermediate scale Quantum)

I. INTRODUCTION

The pharmaceutical industry faces a constant challenge:developing safe and effective drugs in a timely and cost-efficient manner. Traditional drug discovery is a lengthy and expensive process, often riddled with trial and error. However, the emergence of quantum machine learning (QML) presents a potential paradigm shift for the industry .

This introduction will explore the limitations of current drug discovery methods and how QML offers a promising solution for optimizing pharmaceutical data analysis.Here are some key points to be covered in the introduction: Highlight the importance of drug discovery and the challenges faced by the pharmaceutical industry.Briefly explain the traditional drug discovery process and its limitations, such as time-consuming clinical trials and high costs.Introduce the concept of quantum computing and its potential to revolutionize various fields, including drug discovery.Define quantum machine learning (QML) and its unique capabilities, such as harnessing the power of quantum mechanics to accelerate data processing and analysis.Briefly mention the potential applications of QML in optimizing pharmaceutical data, including:Molecular modeling and simulation with better accuracy[7].Efficient selection of promising drug candidatesEnhanced drug safety assessmentsConclude by emphasizing the potential of QML to expedite drug discovery, reduce costs, and ultimately improve human health.This introduction provides a starting point for delving deeper into the exciting world of QML and its potential to revolutionize the pharmaceutical industry.

II. RELATEDWORK

The existing system is used to check the molecular design, information and find the functional groups.In drug discovery identifying hits from compound's historical data that may become leads for medicinal chemistry optimizationtechnique[3]. There are also protein folding simulations studied by pharma companies R&D sections. They might give a picture of reaction between target proteins i.e. based on ligands of the micro-organism causing disease. This cannot alone justify the bioactivity or reaction on the body of patient or other anthropological factors like accepting and rejecting a null hypothesis of medicine.This updated overview discusses the utilization of compound libraries, compounds derived from combinatorial and parallel synthesis campaigns and natural product sources; creation of mother and daughter plates; and compound storage, handling, and bar coding in HTS.The unit also presents an overview of established and emerging assay technologies (i.e., time-resolved fluorescence, fluorescence polarization, fluorescence-correlation spectroscopy, functional whole cell assays, and high-content assays) and their integration in automation hardware and IT systems. There are many recent discoveries in pharma companies. We were dependent on the criterias like hydrogen bond donor, hydrogen bond acceptor, molecular weight and p-IC50 value. Most of the researches in drug discovery are done on the basis of Lipinski's rule of 5 in designing orally consumable drugs.Along with the activity, other factors that determine the efficacy and potency of the hits, such as the absorption, distribution, metabolism, excretion and toxicity (ADMET) profile, among other pharmacokinetic properties, are optimized to produce a smaller set of better.candidates called lead compounds[5]. Author described detailed Survey On Creating Digital Health Ecosystem with Lifewellness Portal Including Hospital and Insurance Company with Cloud Computing and Artificial Intelligence.[12].

III.PROPOSEDSYSTEM

We have achieved quadratic time speed up by Classical to Quantum transfer of information where we can increase the accuracy of the model for further processing.Problems that have time complexity of thousand of year can be solved in a day (only if hardware is noise free).We can conquer any pandemic like 'Corona- SARS V2' or severe disease by new drug design methodology and technology(Quantum Machine Learning) in near future. Those lead compounds again can be refined by newer technique proposed by applying quantum model

which consists of algorithms like QAOA and quantum models. QSAR models for smaller datasets play a major role in the optimization, allowing for quickly judging whether certain modifications improve drug-likeness or not, especially when no target information is available. One thing that is implemented is on which side the risk factor lies. We have designed a metric that is based on generalized score which then is labelled depending on the WHO standardized rules for risk assessment. The Classification criteria given are,Severity(s): How significant the identified risk in terms of impact. Probability (p): How often is frequency of occurence.Detectability (d): includes methods to detect deviations or their associated parameters. The values are mapped by the program related to generalised score considering the terminologies stated. This system considers the clinical trials effectively. Drugs with a level of specific toxicity gets highlighted onto the interface. Correct result is constructively interfered and wrong results are destructively interfered based on quantum gates. This will potentially point out the classified drug name causing more or less harm. This is a highly time-saving in lead optimization introducing a factor in classification named 'activity'.

The bipartite graph i.e. nothing but the Hamiltonian a combinatorial graph formulation of a problem, helps a lot in assessing on which side the risk lies as discussed earlier and also the attributes or co-factors linked with the risk. This potentially also helps pharma companies in figuring out how the selected drugs in the production can or cannot cause the health issue on the overall population. In this way the drug safety profile of pharma companies will get benefited by prohibiting the collection on severe drugs in early production period. With highly fault tolerant Quantum Computers, this process will even get much faster with no time. The quadratic time speed up achieved from superposition and entanglement of bits is quite essential. A breakthrough in hardware and better understandings of complex mathematics behinds quantum mechanics willeliminate the imbalance in qubits stability, making the quantum advantage a true boon to humanity.

IV. QUANTUM APPROXIMATE OPTIMIZATION ALGORITHM:

Quantum Approximate Optimization Algorithm (QAOA) is a hybrid quantum-classical variational algorithm designed to tackle combinatorial optimization problems[8]. Despite its promise fornear-term quantum applications, not much is currently understood about QAOA's performancebeyond its lowest-depth variant. An essential but missing ingredient for understanding and deploying QAOA is a constructive approach to carry out the outer-loop classical optimization.

It provides an in-depth study of the performance of QAOA on MaxCut problems by developing an efficient parameter-optimization procedure and revealing its ability to exploit non-adiabatic operations[8].Building on observed patterns in optimal parameters, we propose heuristic strategies for initializing optimizations to find quasi-optimal p-level QAOA parameters in O(poly(p)) time, whereas thestandard strategy of random initialization requires 2O(p)optimization runs to achieve similar performance

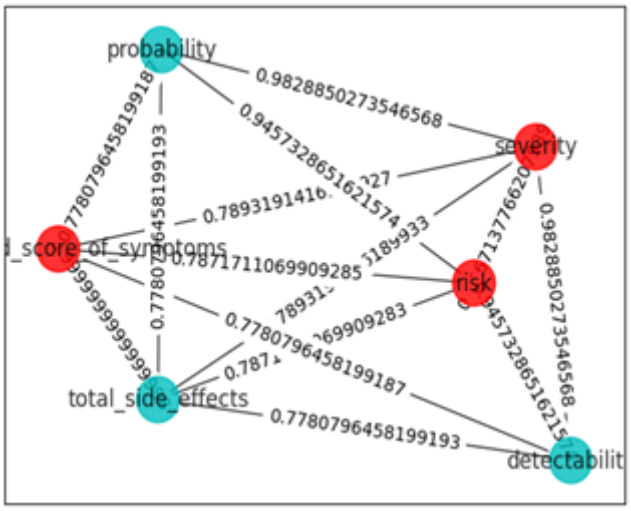

Fig. 1 Optimization Algorithm- QAOA Hamiltonian

We then benchmark QAOA and compare it with quantum annealing, especially on difficultinstances where adiabatic quantum annealing fails due to small spectral gaps. The comparison reveals that QAOA can learn via optimization to utilize non-adiabatic mechanisms to circumvent thechallenges associated with vanishing spectral gaps. Finally, we provide a realistic resource analysison the experimental implementationof QAOA. When quantum fluctuations in measurements areaccounted for, we illustrate that optimization will be important only for problem sizes beyond numerical simulations, but accessible on near-term devices. We propose a feasible implementation oflargeMaxCut problems with a few hundred vertices in a system of 2D neutral atoms, reaching theregime to challenge the best classical algorithms.

A. Quantum Kernel

The idea behind quantum kernels is to take advantage of quantum mechanics to achieve better results while mapping the feature vector[9]. The general guidelines for such mapping are encoding classical data into qubits, performing operations (such as superposition and rotations in the Bloch sphere), then computing the dot product of the resulting states.

Some of the proposals for the kernel have been made. For instance, Aram Harrow et al. proposed the following kernel circuit:

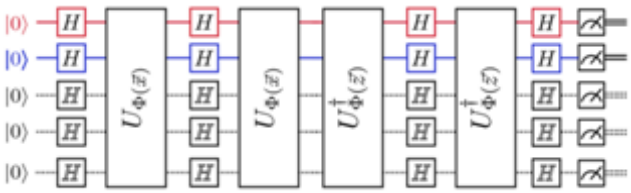

Fig. 2 Quantum Kernel

B. Harrow–Hassidim–Lloyd algorithm.

The Harrow–Hassidim–Lloyd algorithm or HHL algorithm is a quantum algorithm for numerically solving a system of linear equations, designed by Aram Harrow, Avinatan Hassidim, and Seth Lloyd. The algorithm estimates the result of a scalar measurement on the solution vector to a given linear system of equations.Firstly,

the algorithm requires that the matrix displaystyle A} be Hermitian so that it can be converted into a unitary operator. In the case where displaystyle A} is not Hermitian, define

$$C = \begin{bmatrix} 0 & A \\ A\dagger & 0 \end{bmatrix}$$

Although it may be complex, quantum kernels have to perform operations that are not available to a classical machine to outperform them. In the case above, the circuit uses Hadamard gates and Z Pauli matrices to get some advantage over them.

$$U_{\Phi(\vec{x})} = \exp\left(i \sum_{S \subseteq [n]} \phi_S(\vec{x}) \prod_{i \in S} Z_i\right)$$

Unitary gate for a classical feature function ϕ

System Architecture

- Inthismoduleanalysthastologinbyusingvalidusernameandpasswordwhichwerebecross-checkedwith our database. If user is authenticate then user canloginsuccessfullyotherwiselogingetfailed.
- After successfully login the analyst can the upload the dataset of collected drug molecules samples[10].
- Then by simply putting the dataset, the analystcananalyse the scenarios. This will enable him or her classify drugs properly and apply right set of configured drugs.Automatically after entering parameter values hitting categorise button, an event will call a quantum model function that comprises of quantum neural net will make prediction whether new drug is capable of passing the tests compared to other drugs into the cluster.
- Another aspect is KPIs (Key Performance Indicators)provides the sales histogram of current market. Then derivative pricing is predicted for the specific drug. Such decision support system helps in deciding , the introduction of new drug can affect market.
- Information retrieval is the sub-part involved in statistical analysis of the production.
- The newly generated dataset, graphs and reports from researchers are again checked. The changes are updated in the molecules database, policy formulation in drug discovery by board of members or AI assistance regarding drug safety.

Fig.3.Systemarchitecture

C. Figures

1) Activity Diagram

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In UML's Activity diagrams show the overall flow of control is shown. The interaction of various processes and tasks is modelled. Graphical representation: Rounded rectangles represent actions; Diamonds represent decisions; Bars represent the start (split) or end (join) of concurrent activities; A black circle represents the start (initial state) of the workflow. An encircled black circle represents the end (final state). Arrows depict the process flow in activity diagram. From which terminal it originated and where it will head towards the body is denoted by the arrows. Swimlanes shows the partitioning between the modules. How much parallel tasks to be kept, how much concurrency is to be provided. The design tells that no matter how the modules are divided, the control, the flow and behaviour will be definitely mapped in this type of dynamic model i.e. Activity model. It truly says a lot about the real world objects. Here, it typically describes user control, processes, entities and decision making through forks joins and splits and other UML components

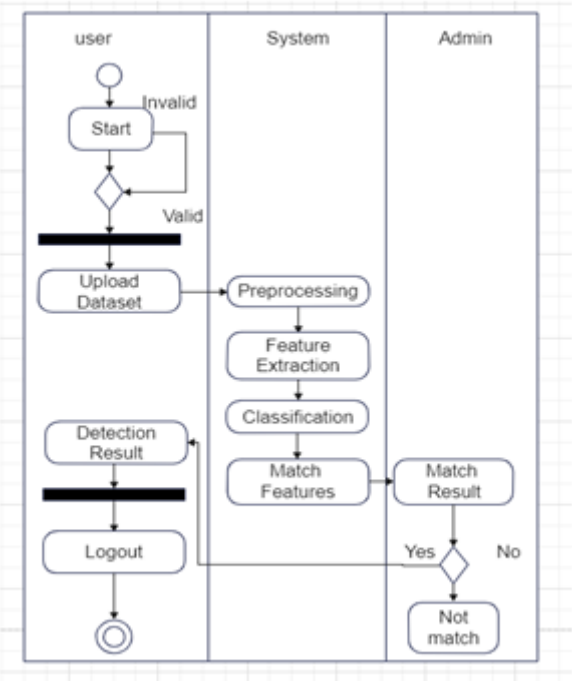

Fig.4.Activity Diagarm

2) DFD Level Diagram

Level 0

Level 1

Fig.5.DFD Level Diagram

3) Deployment Diagram

In this section, we are going to have an overview of Deployment diagrams that is used for visualizing the deployed software components in a toplogical manner. The connection between nodes, their relationship is described. In our project we have a jupyter notebook deployed on IBM server and datasets.This notebook contains most important function that maked prediction. It consists of pre-processing, feature extraction and classification algorithm. The quantum model does classification. It comprises 6 layered quantum neural net

made from quantum gates. Those gates can be runned on a real quantum hardware connecting the real quantum hardware available online through IBMQ account. That notebook is connected to the application server that has other files, custom database (user management) and system libraries that gives access to server side scripts in anvil framework (HTML, CSS, python, bootstrap and python). The deployment diagram when detailed functionalities of Model-view-Control architecture are included, it can then be evolved into a 'Component Diagram'.

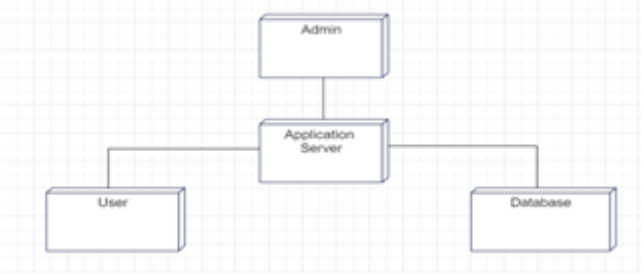

Fig.6.DeploymentDiagram

V. MATHEMATICALMODEL

TheModelingProcedureHasBeenDoneaccordingtotheFollowingAssumptions:
LetSbetheWholesystemS
Input:
Identify the inputs F= f1, f2, f3 ....., FN— F as set of functions to execute commands.

- I= Severity , Probability, Detectability , Activity of Drug
- O= Drug name after classification, Drug Suggestion
- S= Preprocessing, Extraction, Classification, Feture Mapping
- Risk=Severity*Probability*Detectability

Space Complexity:

1. The space complexity depends on Presentation and visualization of discovered patterns.
2. More the storage of data more is the space complexity.
3. Size of the program, generated temporary state space or variable consumes space.
4. Extra generated field after one-hot encoding may increase the size of dataset if we want modify the dataset.

Time Complexity:

1. Check No. of patterns available in the datasets= n
2. If (n(1)) then retrieving of information can be time consuming.
3. So the time complexity of this algorithm is O (nˆn). = Failures and Success conditions.
4. Increasing the number of layers in the quantum model can lead to minute increase in time.

VI. HARDWAREANDSOFTWAREREQUIREMENTS

A. HardwareRequirements

- Processor - Intel i3/i5/i7

- Speed - 3.1 GHz
- RAM - 2GB(min)
- Virtual Machine
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor – SVGA

B. Software Requirements

- Application Server - Apache Tomcat
- Front End - HTML,Bootstarp,CSS
- Language - Python.
- Database -Cloud
- IDE - Pycharm,VisualStudio,IBM Platform

VII. RESULT ANALYSIS

Hereistheresultanalysisforourproposedsystem.

Fig. 7 Login and Registration

Fig.8Information Retrieval- Sales

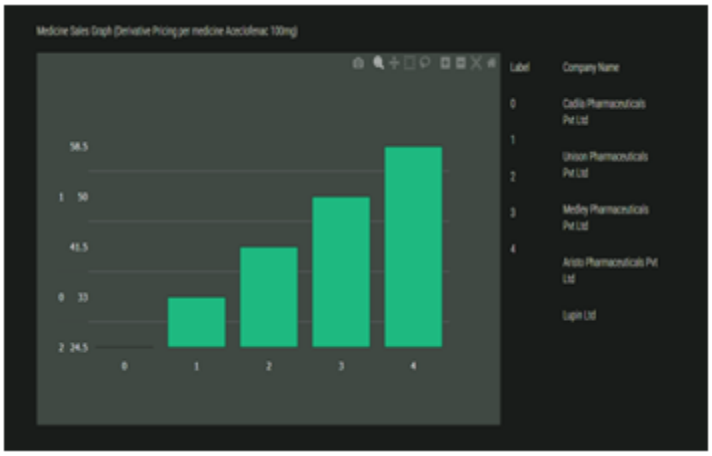

Fig. 9 Histogram- KPI

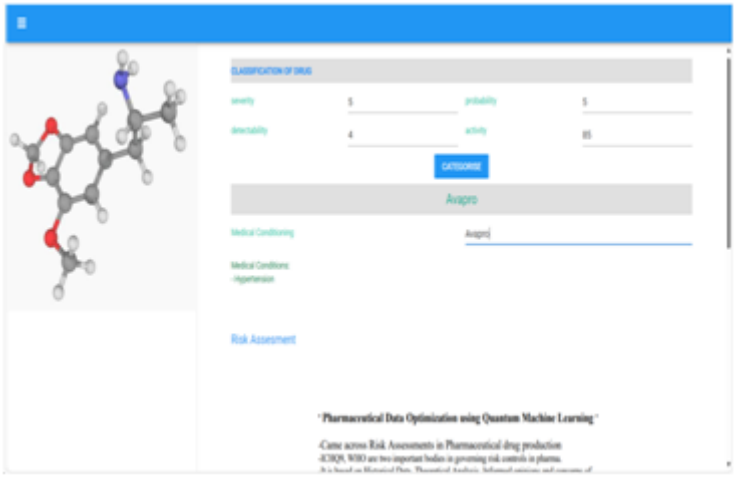

Fig. 10Classification- Quantum machine learning

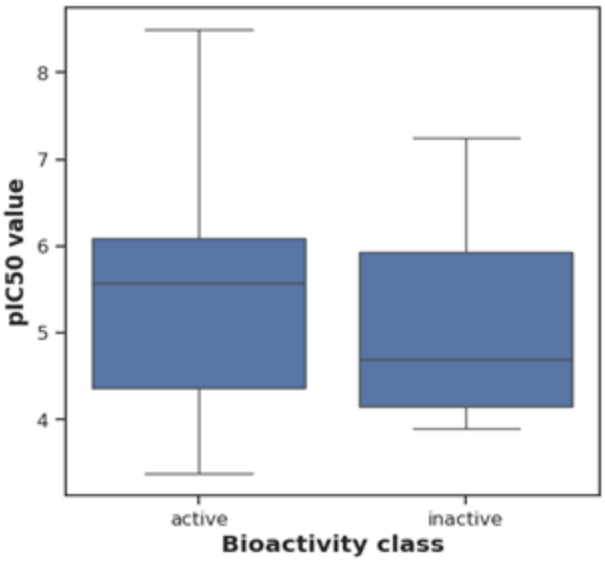

Fig. 11Bioactivity- Selected target protein(prostaglandin)

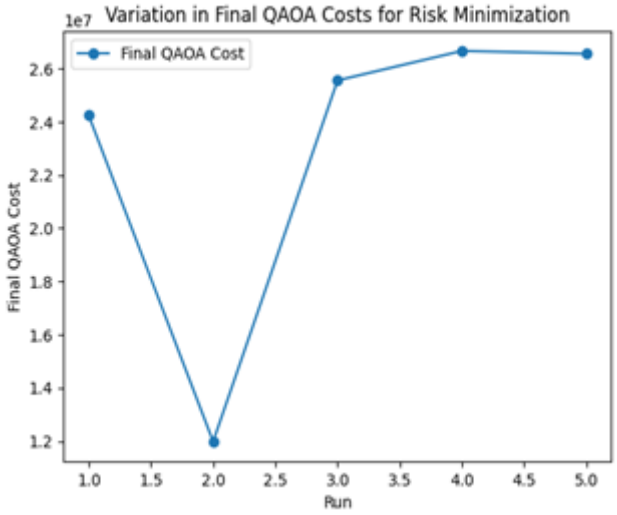

Fig.12Risk assessment- Drug safety production

VIII. COMPARISON

As we compare proposed system with existing system, the proposed system is more faster and efficient. It is due to use of quantum bits in our proposed system. Our proposed system provides lots of advantages like HTS, Lead Optimization, Drug Safety with newer metrics Applied, Quadratic time speed up. It also provides us drug name with medical conditioning.

IX. CONCLUSION

Our project we have accomplished risk assessment according to WHO (World Health Organisation) guidelines and performed QAOA and classification. It is not about proteins structure but an alternative method that is capable, of predicting drug safety, drug testing and business intelligence that can direct towards the drugs that needs enhancement, then indirectly point out drug improvement by combining classical machine learning for drug discovery.

X. ACKNOWLEDGMENT

We would also like to give special thanks to Mr. KarthiganeshDurai, (Chief Quantum Architect, BosonQ Psi) for giving idea about appropriate algorithms to use in the project development. Also greetings to Dr. S.T. Shirkande (Principal at SBPCOE, Indapur), giving opportunity by accepting the idea and Dr.A.B. Gavali for sharing their pearls of wisdom with with us to dot the research regarding the “Pharmaceutical Data Optimization using Quantum Machine Learning”.Also,thanksMr.SangramDeshpande(Senior Quantum Researcher, Qkrishi) for approving our project.

XI. REFERENCES

[1]. 1. Paula Carracedo-Reboredo, Jose Linares-Blanco, NereidaRodr´ıguez-Fern´andez, Francisco Cedron, Francisco J., “A review on machine learning approaches and trends in drug discovery”,IEEE 2021.

[2]. KushalBatra, Kimberley M. Zorn, Daniel H. Foil, EniMinerali, Victor O. Gawriljuk, “Quantum Machine
[3]. JUNDE LI, RASIT O. TOPALOGLU, AND SWAROOP GHOSH, “Quantum Generative Models for Small Molecule Drug Discovery”, IEEE 2018.l
[4]. MaximillianZinner, Florian Dahlhausen, Philip Boehme, Jan Ehlers, Linn Bieske, Leonard Fehring, “Quantum computing’s potential for drug discovery: Early stage industry dynamics”, IEEE 2021.
[5]. Yudong Cao, Jhonathan Romero and Alan Aspuru-Guzik, “Potential of quantum computing for drug discovery”, IEEE 2018..
[6]. TARIQ M. KHAN, AND ANTONIO ROBLES-KELLY, “MachineLearning: Quantum vs Classical”, IEEE 2020.
[7]. Rongxin Xia Sabre Kais, “Quantum machine learning for electronic structure calculations”, IEEE 2018.
[8]. Shree Hari Sureshbabu1, Dylan Herman1, Ruslan Shaydulin1, Joao Basso2, Shouvanik Chakrabarti1, Yue Sun1, and Marco Pistoia1 “Parameter Setting in Quantum Approximate Optimization of WeightedProblems”
[9]. Glick, J.R., Gujarati, T.P., Córcoles, A.D. et al. “Covariant quantum kernels for data with group structure”.Nat. Phys. (2024)
[10]. ChaninNantasenamat, AvratanuBiswas, J.M.Napolesand Mitchell I. Parker, “Building bioinformatics web applications with streamlit”,Inbook: Cheminformatics, QSAR and Machine Learning Applications for Novel Drug Development.
[11]. Aaglave, K. N., ShivanjaliSantoshJadhav, AmaanFirojKhatib, and RohiniLaxmanKhurangale. "A Survey on the Web Scraping: In the Search of Data." (2023).
[12]. Nalawade, V. S., Jadhav, O. D., Jadhav, R. M., Kargal, S. R., &Panhalkar, N. S. (2023). A Survey On Creating Digital Health Ecosystem with Lifewellness Portal Including Hospital and Insurance Company with Cloud Computing and Artificial Intelligence.
[13]. Shirkande, Shrinivas T., et al. "An Explainable Deep Learning Model for Clinical Decision Support in Healthcare." WSN and IoT. CRC Press 225-242.
[14]. Sakhare, N. N., Limkar, S., Mahadik, R. V., Phursule, R., Godbole, A., Shirkande, S. T., &Patange, A. (2023). Ethical considerations of AI applications in medicine: A policy framework for responsible deployment. Journal of Krishna Institute of Medical Sciences (JKIMSU), 12(4).
[15]. Bangare, J. L., Kittad, N., Hundekari, S., Sable, N. P., Shirkande, S. T., &Dhaigude, T. A. (2023). The Intersection of Technology and Public Health: Opportunities and Challenges. South Eastern European Journal of Public Health, 81–96.