Hello everyone, thank you for joining me today. [My name is Yifan Zhou and I’ll be representing our team to present our paper: “The New Answer to Drug Discovery: Quantum Machine Learning in Preclinical Drug Development”](https://link.springer.com/article/10.1007/s10822-020-00346-6) .

our paper proposes a novel concept of using quantum computing and machine learning to improve the drug development process. In this presentation, I will explain the background, the methods, and the potential benefits of this approach.

First, let me give you some background information and current challenges and trends in drug R&D. Durg R&D is where researchers have to identify potential drug candidates, test their efficacy, safety, and optimize their properties. This phase involves a lot of trial and error, experimentation, and simulation, which can be time-consuming and resource-intensive.

As you may know, developing a new drug is a long and costly endeavor, often taking several years and billions of dollars to bring a new drug to market. According to a study by DiMasi et al[2](https://arxiv.org/abs/2201.04093" \t "_blank), the average cost of developing a new drug in 2013 was estimated to be $2.6 billion, which is more than double the cost in 2003. Moreover, the success rate of new drug candidates is very low, with only about 10% of them reaching the final approval stage [3](https://ieeexplore.ieee.org/document/9673630/" \t "_blank).

To overcome these challenges, researchers have been using various computational tools and techniques to assist in drug discovery. [For example, computer-aided drug design (CADD) uses various algorithms and databases to predict the behavior of molecular systems and identify new drug targets 4](https://www.mckinsey.com/industries/life-sciences/our-insights/pharmas-digital-rx-quantum-computing-in-drug-research-and-development). There is also [Molecular dynamics simulations (MDS) which use physics-based models to simulate the interactions and motions of molecules in different environments 5](https://arxiv.org/abs/2104.00746).

However, these methods also have limitations. For instance, CADD relies on existing knowledge and excessive data, which may not capture all the possible molecular interactions and variations. Moreover, MDS requires a lot of computational power and time to simulate complex systems with high accuracy.

This is where our solution comes in. Quantum-based machine learning simulation, or QMLS for short combines the power of quantum computing and machine learning to simulate and predict the behavior of complex molecular systems, such as drug-target interactions, drug metabolism, and drug toxicity. QMLS can potentially offer advantages over classical methods, such as faster computation, higher accuracy, and lower cost.

The basic framework for QMLS consists of three main components: Machine Learning Molecule Generation (MLG), machine learning variation (MLV), and quantum-based simulation (QS). The system works in accordance to drug R&D. For Hit Generation (fist stage…), this phase is parallel to the Machine Learning Molecule Generation (MLG) generates possible hits according to the molecular structure of the target protein while the Quantum Simulation (QS) filters these molecules from the primary essay based on the reaction and binding effectiveness with the target protein. Then, For Lead Optimization (in which…), the resultant molecules generated and filtered from MLG and QS are compared, and molecules that appear as a result of both processes will be made into dozens of molecular variations and optimization through Machine Learning Molecule Variation (MLV). Lastly, all optimized molecules would undergo multiple rounds of QS filtering with a high standard for reaction effectiveness and safety, creating a few dozen pre-clinical-trail-ready drugs.

Basically, QS & QML take s place same time. They correspond to the Hit Generation stage of R&D. The purpose of this is … to accompany both novel creation(ml) and existing solutions (QS). Then results compared if match... wich means … MLV create optimation & variation (leda optimization.)

MLG is the first component of QMLS, which is responsible for generating possible drug candidates based on the molecular structure of the target protein or gene.

**Explain Picture**

**Transfer Learning… advantage-> break down complex task**

Moreover, MLG uses quantum generative models, such as quantum generative adversarial networks (QGANs) or quantum variational autoencoders (QVAEs), to learn the probability distribution of the molecular space and sample new molecules from it.

The advantage of using QGANs for QMLG is that they can generate diverse and novel drug candidates that are not limited by existing data or knowledge. They can also learn complex features and correlations in the molecular space that may be hard to capture by classical methods. Moreover, QGANs can leverage quantum resources to speed up the training and sampling process, as well as enhance the quality and diversity of the generated molecules.

QS is the second component of QMLS, which is responsible for filtering out molecules based on their reaction and binding effectiveness with the target protein or gene. QS performs accurate and efficient simulations of molecular systems. QS use quantum algorithms to simulate the dynamics and interactions of molecules in different environments, such as solvents, temperatures, pressures, etc. Quantum algorithms can offer advantages over classical algorithms in terms of scalability, accuracy, speed, and complexity.

OpenMM can be taken as an example of QS. OpenMM is a software framework for performing molecular simulations on classical and quantum computers. OpenMM allows users to define molecular systems using various force fields, integrators, solvers, etc., and run simulations on different platforms, such as CPUs, GPUs, or quantum computers. OpenMM can also interface with other software packages to perform quantum chemistry calculations on molecular systems.

The advantage of QS is that it can filter out molecules based on their physical and chemical properties, such as energy, stability, reactivity, solubility, toxicity, etc. It can also evaluate the binding affinity and specificity of molecules with the target protein or gene, which are crucial for drug discovery. Moreover, QS can exploit quantum resources to improve the accuracy and efficiency of molecular simulations, as well as explore new regimes of molecular behavior that may be inaccessible by classical methods.

QMLV is the third component of QMLS, which is responsible for creating variations and modifications of the filtered molecules based on their fitness scores. QMLV uses quantum variation models, such as quantum genetic algorithms (QGAs) or quantum random number generators (QRNGs), to optimize the molecular structures and properties according to some objective functions. Quantum variation models are extensions of classical variation models that use quantum circuits to implement mutation and crossover operations (in QGAs) or random sampling operations (in QRNGs). Quantum circuits can introduce more diversity and creativity in the variation process than classical circuits, as they can generate superpositions and entanglements of molecular states.

The authors use QGAs as an example of QMLV in their paper. QGAs are quantum versions of genetic algorithms, which are evolutionary algorithms that mimic the natural selection process. QGAs consist of a population of quantum circuits, each representing a molecular structure. QGAs apply mutation and crossover operations to the quantum circuits to create new molecular structures, and evaluate their fitness scores based on some objective functions, such as binding affinity, stability, solubility, etc. QGAs select the best quantum circuits to form the next generation, and repeat the process until a desired solution is found or a termination criterion is met.

The advantage of using QGAs for QMLV is that they can optimize the molecular structures and properties according to multiple criteria, such as efficacy, safety, bioavailability, etc. They can also generate new molecular structures that are not present in the original population, which may lead to novel and improved drug candidates. Moreover, QGAs can leverage quantum resources to enhance the diversity and quality of the variation process, as well as accelerate the convergence and exploration of the search space.

In summary, QMLS is a theoretical framework that combines quantum computing and machine learning to revolutionize the drug development process. QMLS consists of three main components: QMLG, QS, and QMLV, which work together to generate, filter, and optimize drug candidates based on the molecular structure of the target protein or gene. QMLS has the potential to offer significant advantages over existing methods in terms of efficiency, effectiveness, diversity, novelty, scalability, robustness, applicability etc. However, QMLS is still in its early stages of development and faces many challenges such as qubit constraints on noisy quantum computers; lack of standardized benchmarks; ethical issues; etc. Therefore, more research is needed to validate the feasibility; reliability; validity; generalizability; etc., of QMLS; as well as address its limitations; challenges; risks; etc., before it can be widely adopted and implemented in the real world.

Thank you for your attention. I hope you enjoyed this presentation and learned something new.