EXPLORING QUANTUM HORIZONS, PERFORMANCE ANALYSIS OF ANNEALING BASED ML ON ADMET DATASETS

ABSTRACT

This study investigates the performance of quantum annealing-based machine learning models on ADMET datasets, which play a crucial role in drug discovery and pharmacokinetic analysis. By comparing Quantum Support Vector Machines (QSVM) and Quantum Boosting (QBoost) with classical LightAutoML approaches, the research explores their strengths and limitations in handling chemical descriptors derived from SMILES strings. Experimental results reveal that while classical models currently outperform quantum counterparts in consistency, quantum methods—particularly QBoost—show promising potential in specific contexts, achieving competitive AUC-ROC scores. The findings suggest that the integration of hybrid quantum-classical solvers could redefine computational paradigms for high-dimensional biological data, offering pathways toward enhanced accuracy and scalability in molecular property prediction.

Keywords: Quantum computing, quantum annealing, QSVM, QBoost, ADMET datasets, drug discovery, LightAutoML, hybrid solvers.

EXISTING SYSTEM

The current computational landscape for ADMET prediction primarily relies on classical machine learning frameworks like LightAutoML, Random Forests, and Gradient Boosting models. These systems handle chemical descriptors efficiently and employ ensemble learning techniques for improved accuracy. However, they depend heavily on handcrafted feature extraction and suffer from scalability issues when dealing with exponentially growing molecular datasets.

Moreover, traditional systems often struggle with energy-intensive computations and high latency in large-scale optimization. The reliance on cross-validation and hyperparameter tuning introduces computational bottlenecks. While effective in performance metrics such as AUC-ROC, these methods lack the intrinsic parallelism that quantum systems can provide.

Another limitation lies in representational constraints—classical systems encode molecular structures numerically, missing potential correlations embedded in quantum states. As data

complexity increases, such linear or kernel-based approximations fail to capture deeper quantum-level interactions relevant to drug behavior.

Disadvantages of Existing System

1. Computational Inefficiency: High-dimensional chemical datasets result in slow convergence and heavy resource consumption.
2. Feature Dependence: Classical models require manual feature engineering, limiting adaptability across diverse molecular datasets.
3. Scalability Constraints: Performance degrades when models are expanded to handle complex biological interactions or larger data volumes.

PROPOSED SYSTEM

The proposed system utilizes Quantum Annealing Machine Learning (QAML) frameworks—specifically Quantum Support Vector Machine (QSVM) and QBoost—to address these limitations. Both are implemented using D-Wave's Advantage architecture, leveraging the Pegasus topology for enhanced qubit connectivity and hybrid solvers for partitioned problem execution.

QSVM transforms optimization constraints into QUBO formulations, enabling efficient exploration of high-dimensional feature spaces via quantum tunneling. QBoost, on the other hand, iteratively refines weak classifiers using quantum selection mechanisms, dynamically adapting ensemble weights through annealing optimization.

This hybrid system unites classical pre-processing—like SMILES standardization and molecular embedding—with quantum-level optimization, significantly improving computational throughput and potentially reducing model training time. By integrating both paradigms, the system aims to achieve superior performance across multiple ADMET datasets and enhance the interpretability of pharmacological predictions.

Advantages of Proposed System

1. Enhanced Optimization Speed: Quantum tunneling enables faster convergence to global minima compared to classical gradient-based solvers.

2. Reduced Feature Engineering Burden: QUBO encoding allows direct handling of raw molecular descriptors with minimal preprocessing.
3. Scalability and Adaptability: Hybrid solvers can efficiently handle large biochemical datasets, offering improved computational efficiency and predictive reliability.

SYSTEM REQUIREMENTS

- **H/W System Configuration:-**
- Processor - Pentium –IV
- RAM - 4 GB (min)
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

SOFTWARE REQUIREMENTS:

- **Operating system** **:** Windows 7 Ultimate.
- **Coding Language** **:** Python.
- **Front-End** **:** Python.
- **Back-End** **:** Django-ORM
- **Designing** **:** Html, css, javascript.
- **Data Base** **:** MySQL (WAMP Server).