HQGNN-Tox: Hierarchical Quantum Graph Neural Network for Personalized Drug Toxicity Prediction to Prevent Adverse Drug Reactions on NISQ

Spandan Das - Spandand2011@gmail.com

Research Question
Can Hybrid Quantum-Classical Methods for Machine Learning accurately predict personalized drug-toxicity values in low data regimes to prevent Adverse Drug Reactions?

Abstract
In this proposal, we aim to combat the lack of personalization in low-data regimes of drug toxicity prediction. Drug Toxicity is vastly dependent on metabolic status and other demographic metrics. This makes it increasingly hard to predict drug toxicity for many individuals, leading to a gap in the accuracy of current classical ML models. Our proposal uses a Hybrid Quantum-Classical ML Model, A Hierarchical Quantum Graph Neural Network, which uses NISQ hardware to represent drugs with a final probability measurement to arrive at the drug toxicity. In this research, we establish such a framework specifically for IBM Eagle and similar quantum processors to perform these tasks and attempt to close the personalization gap left by current models.

Research Problem & Introduction
Adverse drug reactions cause over 1.25 million injuries and deaths in the U.S. each year, with studies estimating 45% are preventable through better screening before prescribing. The core issue is that current prediction tools treat all patients the same, ignoring individual genetics, medical history, and drug combinations. This leaves doctors unable to assess risk for patients taking a drug for the first time, where classical machine learning models fail due to a lack of prior data (Pu et al., 2019). As a result, clinicians lack reliable tools to prevent harm in the exact cases where patients are most vulnerable. To address this gap, we propose a hierarchical framework that links drug chemistry, population-level pharmacogenomic data, and individual health records to infer drug toxicity, even with minimal patient history. This is needed to close the personalization gap in drug safety and reduce preventable adverse events.

Literature Review
Previous studies for drug toxicity prediction have relied heavily on classical graph methods. Graph Convolutional Neural Networks like Decagon achieved 69% accuracy in polypharmacy side effect prediction by modeling drug-drug interactions on heterogeneous graphs, but require thousands of known drug pairs and fail for new compounds (Zitnik et al., 2018). An Extremely Randomized Trees approach trained on the TOXNET database reported 72% generalized accuracy, yet precision remained low at 0.25, producing 75% false alarms (Pu et al., 2019). Other methods. Similarly, another Support Vector Machine (SVM) model reached an 80% accuracy (Cavasotto et. al., 2022). These studies show a clear 30-35% gap in accuracy in current methods. Classical methods have clearly been explored extensively over the past several decades, yet progress has remained relatively slow. Despite incremental improvements, these approaches still struggle to address key challenges such as low-data regime prediction and true personalization at the individual level. Most existing models rely heavily on large, curated datasets like TOXNET, which limits their applicability in many scenarios.

Methodology
Quantum Encoding - Quantum encoding maps classical data into Hilbert space and is a critical step in quantum machine learning. Common strategies include basis encoding, amplitude encoding, and angle encoding, each with different qubit requirements and circuit depths (Li et

al., 2024; Schuld & Petruccione, 2018). For drug toxicity prediction on near-term devices, we select angle encoding because it requires only one qubit per feature and produces circuits compatible with NISQ hardware (Sim et al., 2019). We use the Tox21 database to obtain molecular descriptors for each compound. For every drug, we extract and normalize four features: molecular weight, LogP, topological polar surface area, and hydrogen bond donors. Each normalized feature is applied as an RY rotation to a dedicated qubit. This produces a product state that uniquely represents the drug's chemical profile while preserving the ability to compute quantum kernels between compounds. The choice of angle encoding over amplitude encoding avoids the exponential state preparation that would be necessary for molecular datasets exceeding 10,000 compounds.

Hierarchical Graph Neural Networks - With quantum drug embeddings, we assign each drug its own node, with edges weighted by drug similarity of Morgan fingerprints and shared protein targets from DrugBank. Using Tox21, we form Tier 1 as a high-level drug-toxicity graph of 12,000 nodes. Using PharmGKB and DrugBank, we construct Tier 2 as cohort graphs where nodes represent patient subgroups defined by metabolic status, age brackets, and sex, with edges weighted by known pharmacogenomic interaction strengths. Tier 3 encodes an individual user's drug usage history as a personal subgraph containing only drugs present in their personal health records. Information passes through tiers via parameterized quantum message passing (Verdon et al., 2019; Huang et al., 2021). Tier 1 drug states are prepared via angle encoding, then CNOT gates entangle drug qubits with Tier 2 cohort qubits if PharmGKB indicates a drug-gene interaction. Tier 2 states subsequently entangle with Tier 3 patient qubits via CNOT gates. The final Tier 4 readout layer consists of a single ancilla qubit entangled with all active Tier 3 patient qubits. Measuring the final probability readout yields a personalized toxicity risk score.

Hardware Implementation and NISQ Constraints - The proposed 4-tier QGNN runs on 50–127 qubit Noisy Intermediate-Scale Quantum (NISQ) hardware like IBM Eagle. We assign one qubit per feature using angle encoding. Tier 1 uses 4 qubits per drug for molecular weight, logP, polar surface area, and hydrogen bond donors from Tox21. Tier 2 uses 2 qubits per patient cohort for CYP2D6 metabolizer status and group-level ADR rate from PharmGKB. Tier 3 adds N features, depending on the number of active medications, with each qubit flagging whether a specific drug is present in the patient's history. Tier 4 uses 1 ancilla qubit for final measurement. Assuming N=5 active medications, the full circuit uses 12 qubits total, which fits within the coherence limits of current hardware (Kim et al., 2023). We train only the inter-tier coupling strengths with the parameter-shift rule, keeping drug features fixed to reduce parameters (Mitarai et al., 2018). This architecture satisfies NISQ constraints while remaining scalable.

Conclusion

We establish a theoretical Hierarchical Quantum Graph Neural Network that maps Tox21 molecular data through pharmacogenomic cohorts to individual patient histories using entangled quantum tiers. By encoding chemical, demographic, and personal drug data in quantum states, the model addresses cold-start personalization and low-data regime prediction, two frontier issues that limit classical toxicity models. The architecture is constrained to a minimum of 12 qubits NISQ hardware like IBM Eagle (Sun et. al., 2025). In future work, we plan to implement the 4-tier circuit on IBM Eagle processors and benchmark personalized toxicity predictions against classical baselines using PharmGKB patient subsets.

Bibliography (APA 7th)

- Huang, R., et al. (2016). Tox21Challenge to Build Predictive Models of Nuclear Receptor and Stress Response Pathways as Mediated by Exposure to Environmental Chemicals and Drugs. *Frontiers in Environmental Science*, 3, 85. https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2015.00085/full
- Huang, H. Y., et al. (2021). Power of data in quantum machine learning. *Nature Communications*, 12, 2631. https://www.nature.com/articles/s41467-021-22539-9
- Kim, Y., et al. (2023). Evidence for the utility of quantum computing before fault tolerance. *Nature*, 618, 500-505. https://www.nature.com/articles/s41586-023-06096-3
- Li, Y., et al. (2024). Comparing Quantum Encoding Techniques. *arXiv:2410.09121*. https://arxiv.org/html/2410.09121v1
- Mitarai, K., et al. (2018). Quantum circuit learning. *Physical Review A*, 98, 032309. https://journals.aps.org/pra/abstract/10.1103/PhysRevA.98.032309
- Pu, L. et al. (2019). eToxPred: a machine learning-based approach to estimate the toxicity of drug candidates. *BMC Pharmacology and Toxicology*, 20, 2. https://link.springer.com/article/10.1186/s40360-018-0282-6
- Schuld, M., & Petruccione, F. (2018). *Supervised learning with quantum computers*. Springer. https://link.springer.com/book/10.1007/978-3-319-96424-9
- Sim, S., et al. (2019). Expressibility and entangling capability of parameterized quantum circuits. *Advanced Quantum Technologies*, 2, 1900070. https://arxiv.org/abs/1905.10876
- Verdon, G., et al. (2019). Quantum Graph Neural Networks. *arXiv:1909.12264*. https://arxiv.org/abs/1909.12264
- Zitnik, M., et al. (2018). Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics*, 34, i457-i466. https://academic.oup.com/bioinformatics/article/34/13/i457/5045770
- *Sun, J., et al. (2025).* Benchmarking Quantum Computers via Protocols: Comparing IBM's Heron vs IBM's Eagle. *arXiv:2503.04377*. https://arxiv.org/abs/2503.04377
- *Cavasotto, C. N., & Scardino, V. (2022).* Machine Learning Toxicity Prediction: Latest Advances by Toxicity End Point. *ACS Omega, 7(51), 47536–47546.* https://pmc.ncbi.nlm.nih.gov/articles/PMC9798519/

Appendix

- Made in LaTeX Overleaf