Integrating Quantum Synergistics and Machine Learning: A Transformative Approach to Drug Discovery

Krishna Sayori Deb, Shail Shah

Stevens Institute of Technology

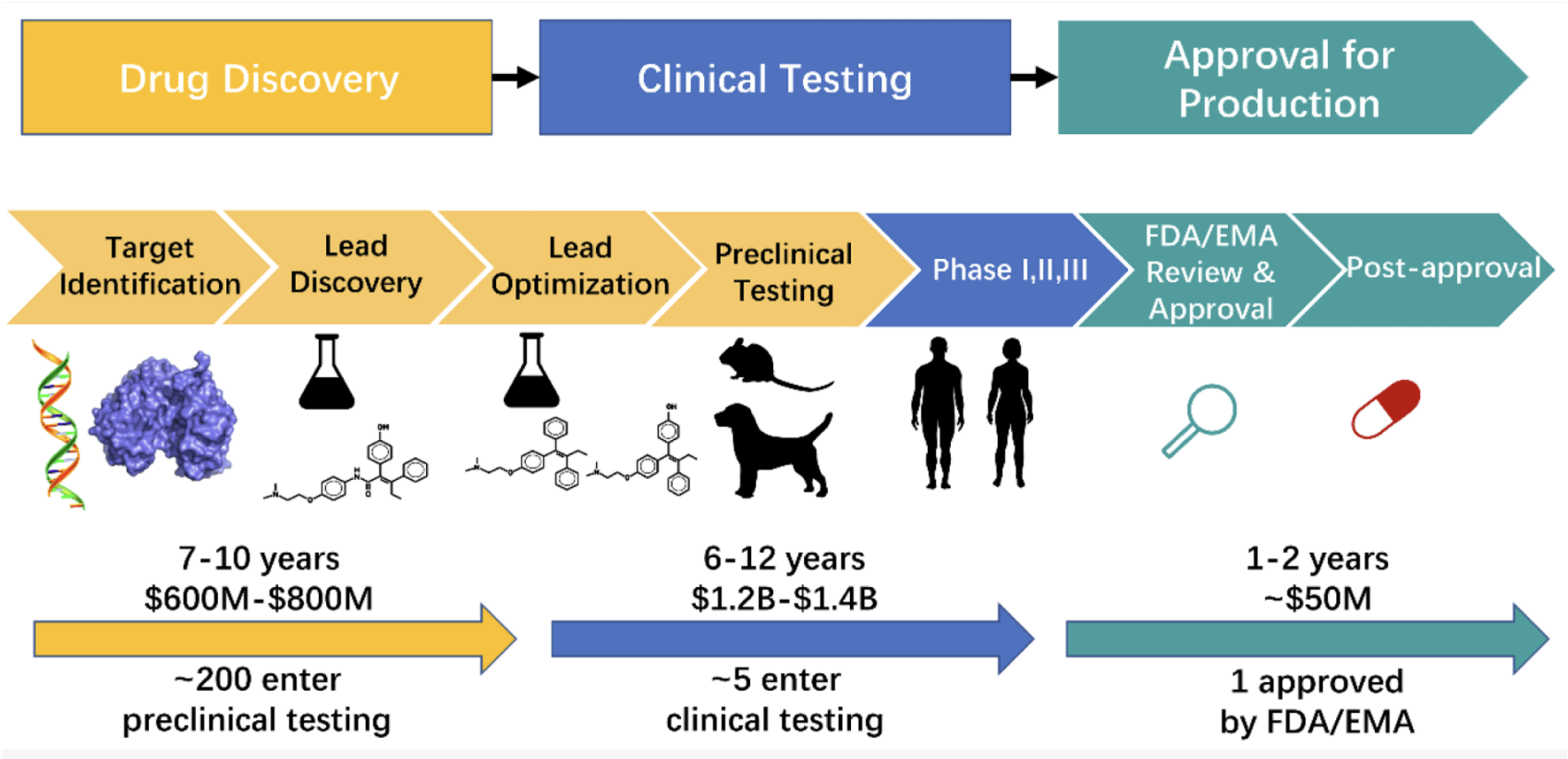
2024

Table of Contents

Table of Contents

Drug Discovery

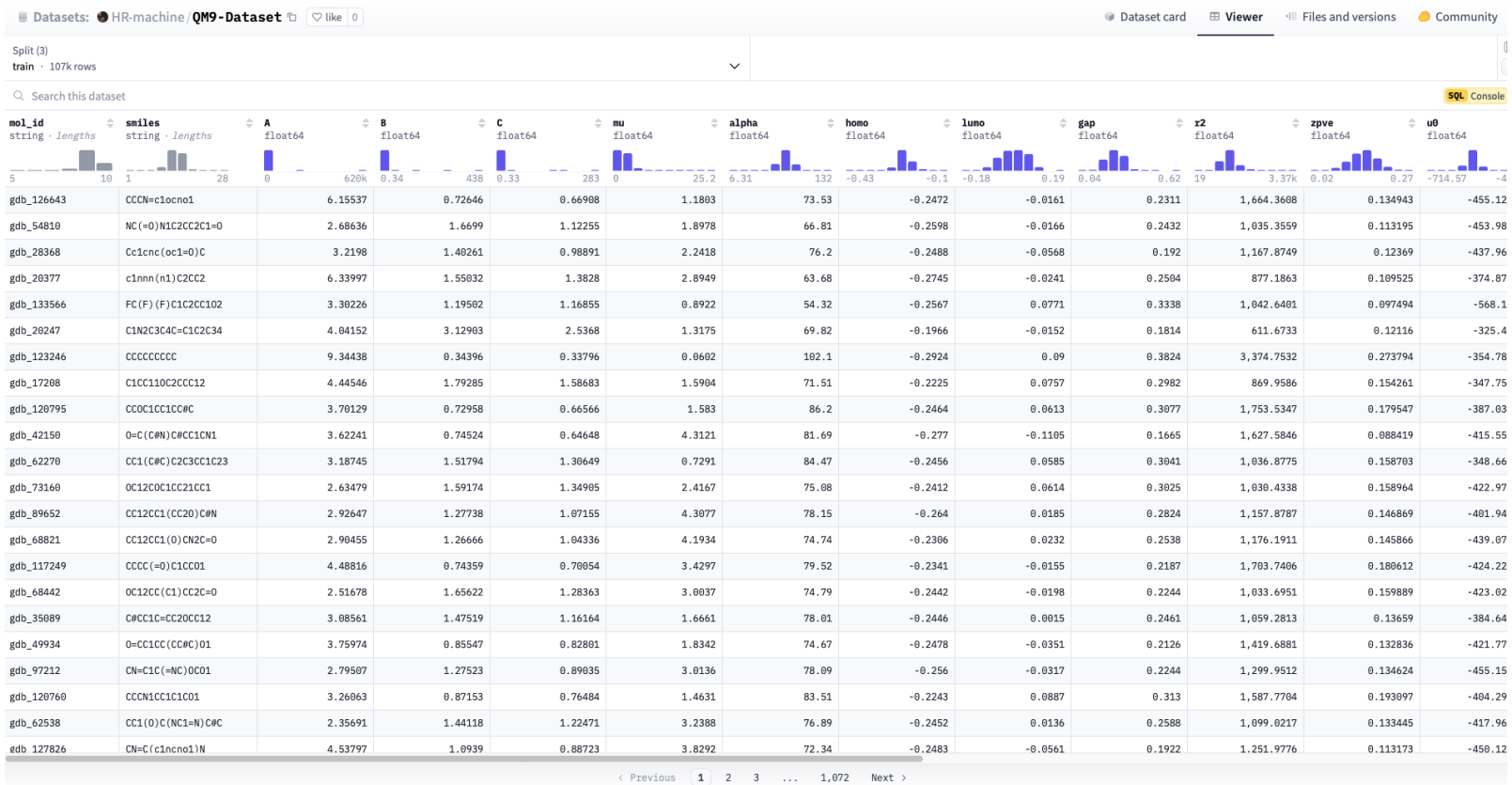
How do we design a drug?



[1]

[1] Y. Zhang et al. (2022). "Application of Computational Biology and Artificial Intelligence in Drug Design". In: *International Journal of Molecular Sciences*.

QM9 Dataset



mol_id	smiles	A	B	C	mu	alpha	homo	lumo	gap	r2	zpve	u0
string · lengths	string · lengths	float64	float64	float64	float64	float64	float64	float64	float64	float64	float64	float64
5 10	1 28	0 620k	0.34 438	0.33 283	0 25.2	6.31 132	-0.43 -0.1	-0.18 0.19	0.04 0.62	19 3.37k	0.02 0.27	-714.57 -4
gdb_126643	CCCN=c1ocno1	6.15537	0.72646	0.66908	1.1803	73.53	-0.2472	-0.0161	0.2311	1,664.3608	0.134943	-455.12
gdb_54810	NC(=O)N1C2CC2C1=O	2.68636	1.6699	1.12255	1.8978	66.81	-0.2598	-0.0166	0.2432	1,035.3559	0.113195	-453.98
gdb_28368	Cc1cnc(oc1=O)C	3.2198	1.40261	0.98891	2.2418	76.2	-0.2488	-0.0568	0.192	1,167.8749	0.12369	-437.96
gdb_20377	c1nnn(n1)C2CC2	6.33997	1.55032	1.3828	2.8949	63.68	-0.2745	-0.0241	0.2504	877.1863	0.109525	-374.87
gdb_133566	FC(F)(F)C1C2CC1O2	3.30226	1.19502	1.16855	0.8922	54.32	-0.2567	0.0771	0.3338	1,042.6401	0.097494	-568.1
gdb_20247	C1N2C3C4C=C1C2C34	4.04152	3.12903	2.5368	1.3175	69.82	-0.1966	-0.0152	0.1814	611.6733	0.12116	-325.4
gdb_123246	CCCCCCCCC	9.34438	0.34396	0.33796	0.0602	102.1	-0.2924	0.09	0.3824	3,374.7532	0.273794	-354.78
gdb_17208	C1CC11OC2CCC12	4.44546	1.79285	1.58683	1.5904	71.51	-0.2225	0.0757	0.2982	869.9586	0.154261	-347.75
gdb_120795	CCOC1CC1CC#C	3.70129	0.72958	0.66566	1.583	86.2	-0.2464	0.0613	0.3077	1,753.5347	0.179547	-387.03
gdb_42150	O=C(C#N)C#CC1CN1	3.62241	0.74524	0.64648	4.3121	81.69	-0.277	-0.1105	0.1665	1,627.5846	0.088419	-415.55
gdb_62270	CC1(C#C)C2C3CC1C23	3.18745	1.51794	1.30649	0.7291	84.47	-0.2456	0.0585	0.3041	1,036.8775	0.158703	-348.66
gdb_73160	OC12COC1CC21CC1	2.63479	1.59174	1.34905	2.4167	75.08	-0.2412	0.0614	0.3025	1,030.4338	0.158964	-422.97
gdb_89652	CC12CC1(CC2O)C#N	2.92647	1.27738	1.07155	4.3077	78.15	-0.264	0.0185	0.2824	1,157.8787	0.146869	-401.94
gdb_68821	CC12CC1(O)CN2C=O	2.90455	1.26666	1.04336	4.1934	74.74	-0.2306	0.0232	0.2538	1,176.1911	0.145866	-439.07
gdb_117249	CCCC(=O)C1CCO1	4.48816	0.74359	0.70054	3.4297	79.52	-0.2341	-0.0155	0.2187	1,703.7406	0.180612	-424.22
gdb_68442	OC12CC(C1)CC2C=O	2.51678	1.65622	1.28363	3.0037	74.79	-0.2442	-0.0198	0.2244	1,033.6951	0.159889	-423.02
gdb_35089	C#CC1C=CC2OCC12	3.08561	1.47519	1.16164	1.6661	78.01	-0.2446	0.0015	0.2461	1,059.2813	0.13659	-384.64
gdb_49934	O=CC1CC(CC#C)O1	3.75974	0.85547	0.82801	1.8342	74.67	-0.2478	-0.0351	0.2126	1,419.6881	0.132836	-421.77
gdb_97212	CN=C1C(=NC)OCO1	2.79507	1.27523	0.89035	3.0136	78.09	-0.256	-0.0317	0.2244	1,299.9512	0.134624	-455.15
gdb_120760	CCCN1CC1C1CO1	3.26063	0.87153	0.76484	1.4631	83.51	-0.2243	0.0887	0.313	1,587.7704	0.193097	-404.29
gdb_62538	CC1(O)C(NC1=N)C#C	2.35691	1.44118	1.22471	3.2388	76.89	-0.2452	0.0136	0.2588	1,099.0217	0.133445	-417.96
gdb_127826	CN=C(c1ncno1)N	4.53797	1.0939	0.88723	3.8292	72.34	-0.2483	-0.0561	0.1922	1,251.9776	0.113173	-450.12

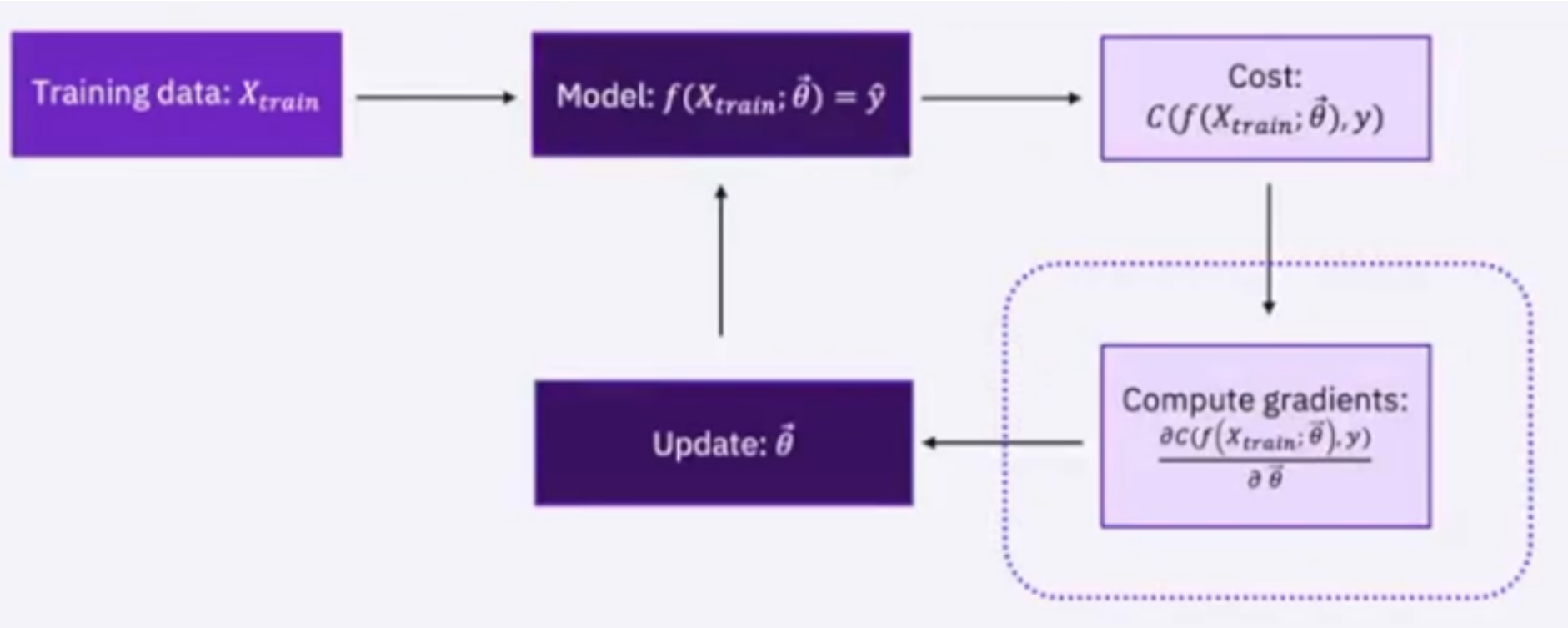
Figure: Around 130k Data points, Here, only HOMO and LUMO value prediction tasks are considered

2

[2] **https://huggingface.co/datasets/HR-machine/QM9-Dataset/viewer**

Table of Contents

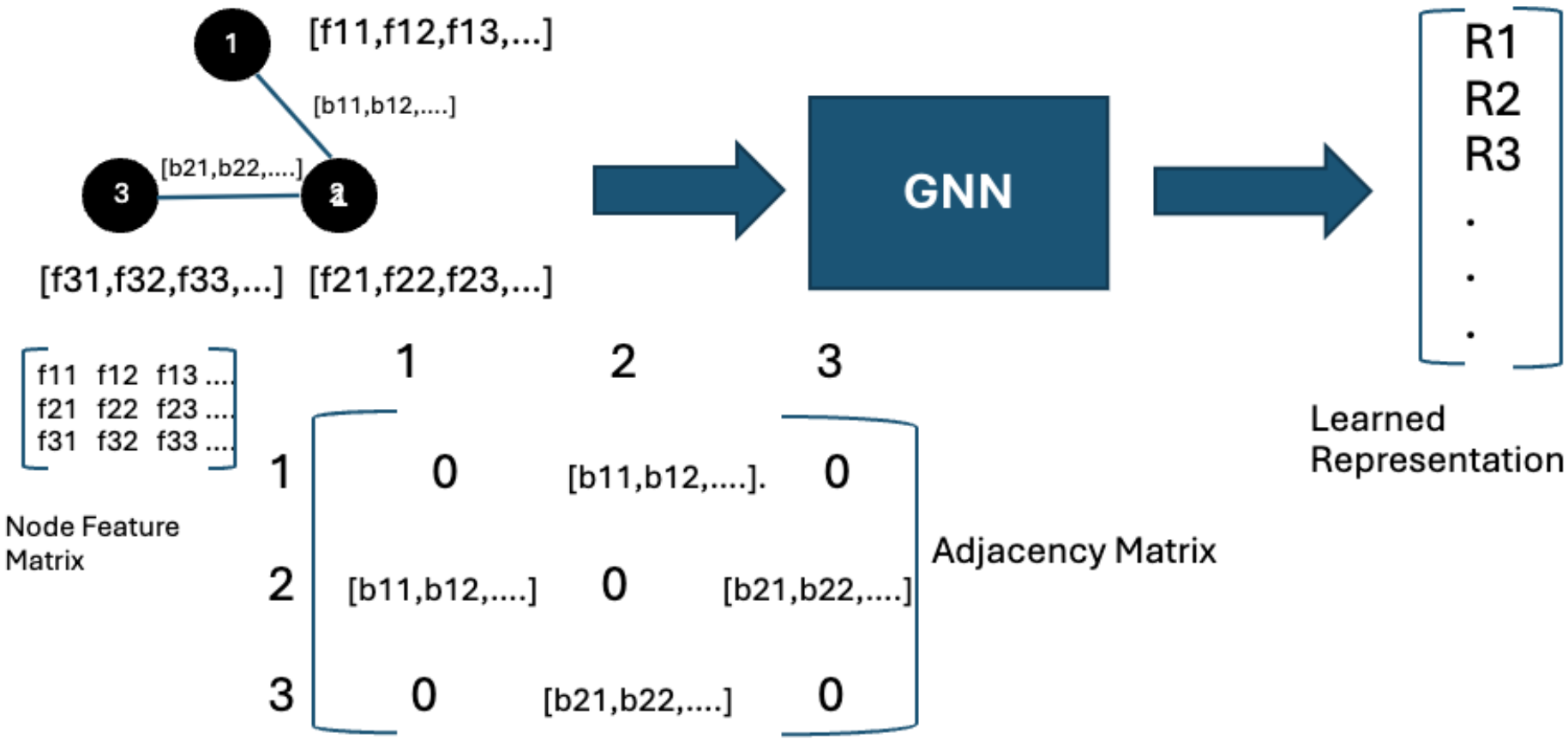
Machine Learning Framework



3

[3] https://www.youtube.com/watch?v=-sxlXNz7ZxU&list=PLOFEBzvs-VvqJwybFx

Graph Neural Networks

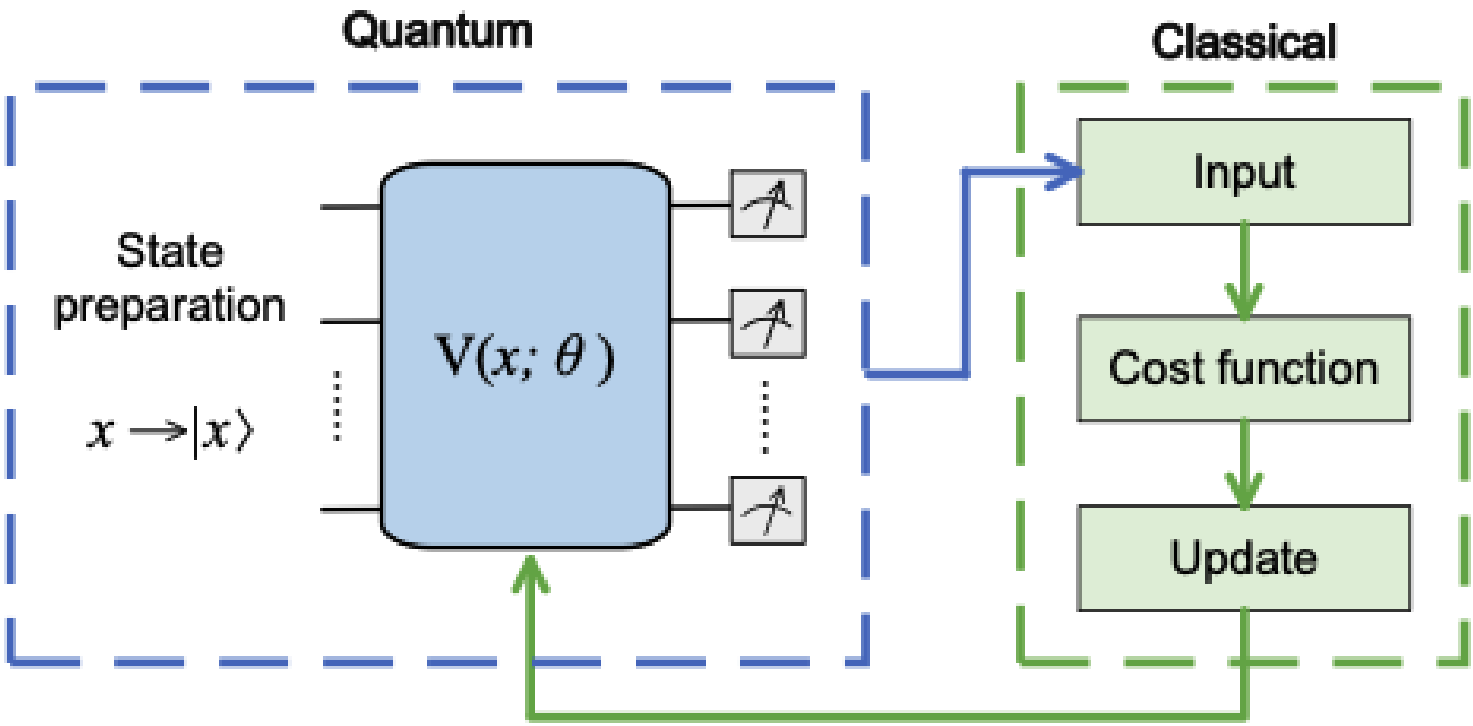


Motivation for Quantum

- **RQ1:** Can quantum computers accelerate the exploration of the vast molecular space in drug discovery?
- **RQ2:** Do quantum computers, leveraging qubits and quantum mechanical phenomena, offer any advantages in learning from scientific data compared to classical machine learning models?

Table of Contents

Variational Quantum Circuits



4

[4] Runqiu Shu et al. (2023). "Variational Quantum Circuits Enhanced Generative Adversarial Network". In: *arXiv preprint arXiv:2308.11062*.

Quantum Data Encoding

There are 3 Encoding Methods to encode classical data points into quantum data[5]

- **Basic Encoding :** Each binary string x_m is mapped to a computational basis state of N qubits. For instance, the binary string (1, 0, 1) would be mapped to the quantum state $|101\rangle$
- **Angle Encoding :**
  - Different from basis encoding, which is limited to binary data. Angle encoding can be used to encode real, floating-point numbers.
  - Data points are then used to rotate the state of a qubit around a specific axis on the Bloch sphere by an angle corresponding to the data. This is done by using rotation gates like Rx, Ry, and Rz, which rotate the qubit state around the X, Y, and Z axes, respectively.
- **Amplitude Encoding :** Represent classical data as quantum states by encoding the classical data into the amplitudes of those quantum states.

[5] Smaldone et al. 2024.

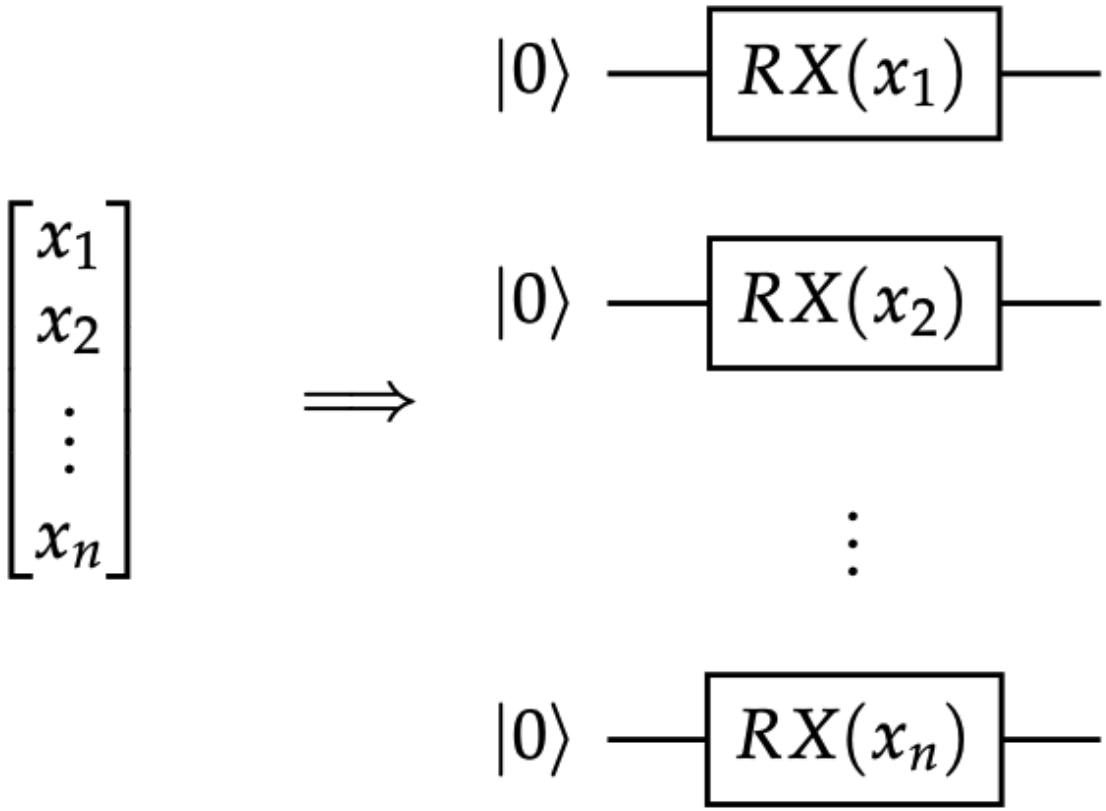


Figure: An example of quantum encoding. Individual elements of an input classical vector is used as the rotation angles of rotation Pauli X (RX) gates, creating a quantum state

6

[6] Ju-Young Ryu, Eyuel Elala, and June-Koo Kevin Rhee (2023). "Quantum Graph Neural Network Models for Materials Search". In: *arXiv preprint arXiv:2308.11759*.

Graph Features

	Feature	Explanation → Data Type
Nodes → Non-Hydrogen Atoms	Atomic Number	C/N/O/F → Integer
	Number of Bonded Hs	0~4 → Integer
	Aromaticity	True/False → Boolean
	Hybridization	sp, sp^2, sp^3 → Integer
Links → Bonds	Bond Type	Single/Aromatic/Double/Triple → Integer

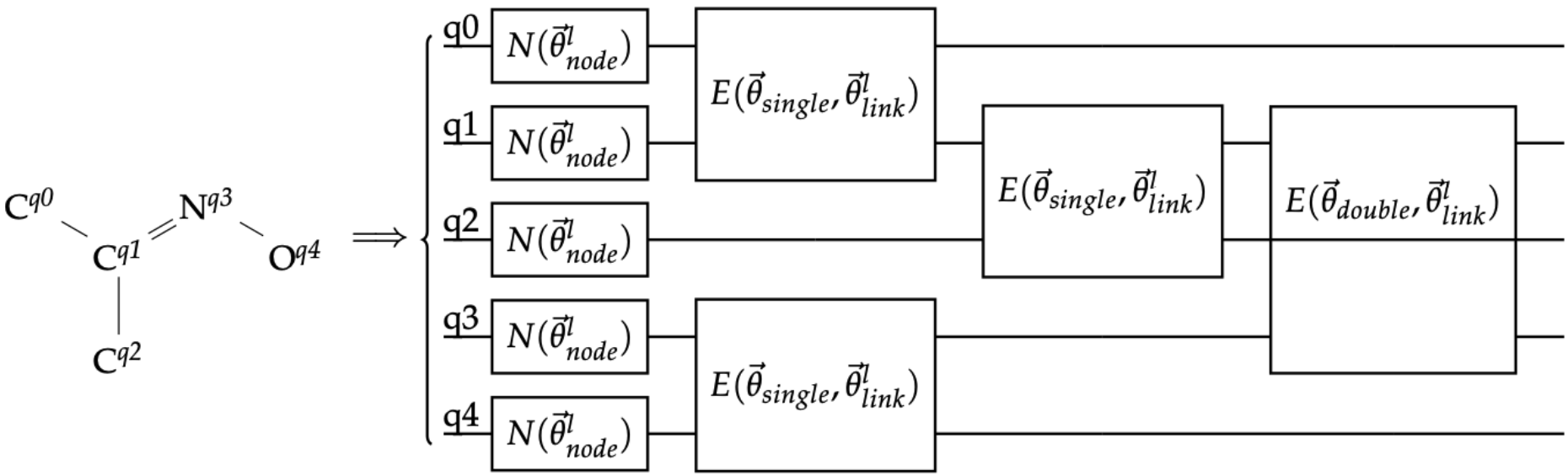
	Atomic Number	Atomic Number and Number of Hydrogens	Atomic Number, Aromaticity, and Hybridization
RY Rotation Angle	$z = 4 \rightarrow 0$, $z = 5,6,7 \rightarrow cos(-1/3)$	$\frac{(2z-7)\pi}{4}$	$\frac{(2z-7)\pi}{4}$
RZ Rotation Angle	$z = 4,5 \rightarrow 0$, $z = 6 \rightarrow 2\pi/3$, $z = 7 \rightarrow -2\pi/3$	$\frac{2\pi n_h}{5}$	$\frac{(-1)^a(2h-1)\pi}{6}$

Figure: Quantum encoding methods. z is the atomic number; nh is the number of bonded hydrogen atoms; a is the aromaticity as True(1)/False(0); h is the hybridization type (sp: 1, sp2: 2, sp3: 3).

7

[7] Ju-Young Ryu, Eyuel Elala, and June-Koo Kevin Rhee (2023). "Quantum Graph Neural Network Models for Materials Search". In: *arXiv preprint arXiv:2308.11759*.

Quantum Graph Neural Network



8

[8] Ju-Young Ryu, Eyuel Elala, and June-Koo Kevin Rhee (2023). "Quantum Graph Neural Network Models for Materials Search". In: *arXiv preprint arXiv:2308.11759*.

Table of Contents

Results

Results from paper[9]

- **Performance:** Improved accuracy over classical GNNs with similar parameters.
- **Efficiency:** Faster convergence during training, leveraging quantum advantages.
- **Adaptability:** Effective handling of diverse molecular structures, scalable to complex datasets.
- **Robustness:** Lower test loss metrics, indicating higher prediction reliability.

[9] Ju-Young Ryu, Eyuel Elala, and June-Koo Kevin Rhee (2023). "Quantum Graph Neural Network Models for Materials Search". In: *arXiv preprint arXiv:2308.11759*.

Comparison Compilation

Classical vs Quantum GNNs[10]

- **HOMO-LUMO Gap Prediction:**
  - **Classical MAE:** HOMO = 0.99, LUMO = 0.87 (enn-s2s Gilmer et al. 2017).
  - **Quantum MAE:**
    - Atomic Number: 1.9975×10^{-2} Hartree.
    - Atomic Number & Hydrogens: 1.8793×10^{-2} Hartree.
    - Aromaticity & Hybridization: 2.0811×10^{-2} Hartree.
- **Hybrid QGNNs:** Competitive with classical models for protein-ligand binding affinity prediction.
- **Quantum CNNs:** Promising results for drug toxicity prediction with reduced training times.

[10] Anthony M. Smaldone et al. (2024). "Quantum Machine Learning in Drug Discovery: Applications in Academia and Pharmaceutical Industries". In: *arXiv preprint arXiv:2405.08294*.

Table of Contents

References

- Gilmer, Justin et al. (2017). "Neural Message Passing for Quantum Chemistry". In: *Proceedings of the 34th International Conference on Machine Learning*. Vol. 70. PMLR, pp. 1263–1272.
- Ryu, Ju-Young, Eyuel Elala, and June-Koo Kevin Rhee (2023). "Quantum Graph Neural Network Models for Materials Search". In: *arXiv preprint arXiv:2308.11759*.
- Shu, Runqiu et al. (2023). "Variational Quantum Circuits Enhanced Generative Adversarial Network". In: *arXiv preprint arXiv:2308.11062*.
- Smaldone, Anthony M. et al. (2024). "Quantum Machine Learning in Drug Discovery: Applications in Academia and Pharmaceutical Industries". In: *arXiv preprint arXiv:2405.08294*.
- Zhang, Y. et al. (2022). "Application of Computational Biology and Artificial Intelligence in Drug Design". In: *International Journal of Molecular Sciences*.