Appendix B: Applied hyperparameters and optimized structures of the classical

neural network for the quantum, classical, and quantum-pretrained classical embed-

ding method

TABLE B1. Applied hyperparameters for the quantum embedding method (NQE with the ZZ feature map)

Quantum embedding	Dataset (class label)	Batch size	Iteration	Learning rate	Binary classifier	Batch size	Iteration	Learning rate
NQE with the ZZ feature map	GBA (1:1 ratio)	25	2000	0.001	QCNN	25	2000	0.05
	KAT2A (1:1 ratio)	25	2000	0.001		128	2000	0.06
	ESR1 antagonist (1:1 ratio)	25	2000	0.001		25	2000	0.01
	MAPK1 (1:1 ratio)	25	2000	0.001		25	2000	0.01
	FEN1 (1:1 ratio)	64	2000	0.001		25	2000	0.007
	PKM2 (1:1 ratio)	64	2000	0.001		25	2000	0.01
	VDR (1:1 ratio)	64	2000	0.001		25	2000	0.009
	ALDH1 (1:1 ratio)	128	2000	0.006		32	2000	0.005
NQE with the ZZ feautre map	GBA (1:6 ratio)	64	2000	0.002	QCNN	32	2000	0.001
	KAT2A (1:6 ratio)	64	2000	0.002		32	2000	0.0005
	ESR1 antagonist (1:6 ratio)	64	2000	0.002		32	2000	0.005
	MAPK1 (1:6 ratio)	64	2000	0.002		25	2000	0.0005
	FEN1 (1:6 ratio)	64	2000	0.001		32	2000	0.0005
	PKM2 (1:6 ratio)	64	2000	0.001		32	2000	0.005
	VDR (1:6 ratio)	64	2000	0.001		32	2000	0.005
	ALDH1 (1:6 ratio)	64	2000	0.001		32	2000	0.005

TABLE B2. Applied hyperparameters for the quantum embedding method (NQE with the XYZ feature map)

Quantum embedding	Dataset (class label)	Batch size	Iteration	Learning rate	Binary classifier	Batch size	Iteration	Learning rate
NQE with the XYZ feature map	GBA (1:1 ratio)	25	1000	0.01	QCNN	30	2000	0.01
	KAT2A (1:1 ratio)	128	2000	0.01		25	2000	0.01
	ESR1 antagonist (1:1 ratio)	64	2000	0.009		20	2000	0.01
	MAPK1 (1:1 ratio)	128	2000	0.009		32	2000	0.01
	FEN1 (1:1 ratio)	128	2000	0.005		20	2000	0.01
	PKM2 (1:1 ratio)	100	2000	0.01		30	2000	0.01
	VDR (1:1 ratio)	100	1000	0.01		32	2000	0.01
	ALDH1 (1:1 ratio)	150	1000	0.01		20	2000	0.01
NQE with the XYZ feature map	GBA (1:6 ratio)	128	1000	0.009	QCNN	64	2000	0.01
	KAT2A (1:6 ratio)	120	1000	0.01		20	2000	0.009
	ESR1 antagonist (1:6 ratio)	256	1000	0.0094		25	2000	0.01
	MAPK1 (1:6 ratio)	256	1000	0.009		32	2000	0.01
	FEN1 (1:6 ratio)	256	1000	0.01		16	2000	0.0095
	PKM2 (1:6 ratio)	128	1000	0.009		30	2000	0.01
	VDR (1:6 ratio)	256	1000	0.0098		64	2000	0.01
	ALDH1 (1:6 ratio)	512	1000	0.01		20	2000	0.01

TABLE B3. Applied hyperparameters for the classical embedding method (the classical counterpart of NQE with the ZZ feature map)

Dataset (class label)	Classical kernel	Batch size	Iteration	Learning rate	Binary classifier	Batch size	Iteration	Learning rate
GBA (1:1 ratio)	RBF	80	1000	1e-7	Single-layer network	100	1000	1e-5
KAT2A (1:1 ratio)		50	1000	1e-6		7	1000	1e-4
ESR1 antagonist (1:1 ratio)		32	1000	1e-4		32	1000	1e-6
MAPK1 (1:1 ratio)		64	1000	8e-4		16	1000	1e-8
FEN1 (1:1 ratio)		32	1000	6e-4		35	1000	1e-5
PKM2 (1:1 ratio)		16	1000	1e-4		50	1000	1e-8
VDR (1:1 ratio)		64	1000	1e-6		50	1000	1e-8
ALDH1 (1:1 ratio)		128	1000	1e-4		1024	1000	1e-4
GBA (1:6 ratio)	RBF	64	1000	1e-3	Single-layer network	10	1000	1e-4
KAT2A (1:6 ratio)		100	1000	1e-6		32	1000	5e-4
ESR1 antagonist (1:6 ratio)		50	1000	1e-4		15	1000	1e-5
MAPK1 (1:6 ratio)		50	1000	1.5e-4		20	1000	1e-6
FEN1 (1:6 ratio)		50	1000	3e-4		30	1000	1e-6
PKM2 (1:6 ratio)		90	1000	1e-6		256	1000	1e-8
VDR (1:6 ratio)		50	1000	1e-7		512	1000	1e-5
ALDH1 (1:6 ratio)		512	1000	1e-5		256	1000	1e-6

TABLE B4. Applied hyperparameters for the classical embedding method (the classical counterpart of NQE with the XYZ feature map)

Dataset (class label)	Classical kernel	Batch size	Iteration	Learning rate	Binary classifier	Batch size	Iteration	Learning rate
GBA (1:1 ratio)	RBF	25	1000	1e-2	Single-layer network	25	1000	1e-7
KAT2A (1:1 ratio)		50	1000	5e-3		25	1000	1e-5
ESR1 antagonist (1:1 ratio)		64	1000	1e-3		40	1000	5e-5
MAPK1 (1:1 ratio)		50	2000	5e-5		35	1000	1e-4
FEN1 (1:1 ratio)		100	1000	8e-5		40	1000	3.5e-4
PKM2 (1:1 ratio)		10	2000	1e-3		30	1000	2e-4
VDR (1:1 ratio)		25	1000	1e-3		27	1000	1e-3
ALDH1 (1:1 ratio)		115	1000	5e-4		128	1000	1e-5
GBA (1:6 ratio)	RBF	10	1000	5e-4	Single-layer network	15	1000	1e-5
KAT2A (1:6 ratio)		64	1000	1e-3		30	1000	1e-5
ESR1 antagonist (1:6 ratio)		25	1000	8e-4		40	1000	8e-6
MAPK1 (1:6 ratio)		50	1000	1e-5		100	1000	1e-8
FEN1 (1:6 ratio)		64	1000	1.5e-4		50	1000	1e-5
PKM2 (1:6 ratio)		50	1000	2e-4		25	1000	1e-5
VDR (1:6 ratio)		64	1000	5e-4		32	1000	5e-5
ALDH1 (1:6 ratio)		100	1000	1e-5		256	1000	1e-5

TABLE B5. Applied hyperparameters for the quantum-pretrained embedding condition 1 (with the classical neural network from the ZZ feature map NQE)

Dataset (class label)	Embedding step	Batch size	Iteration	Learning rate	Embedding step	Batch size	Iteration	Learning rate
GBA (1:1 ratio)	Fine-tuning with RBF	25	2000	1.5×10^{-3}	Classification with the single-layer network	128	2000	7×10^{-13}
KAT2A (1:1 ratio)		200	2000	1×10^{-11}		128	2000	1×10^{-8}
ESR1 antagonist (1:1 ratio)		28	2000	5.5×10^{-4}		8	2000	1×10^{-5}
MAPK1 (1:1 ratio)		128	1000	1×10^{-7}		55	2000	1×10^{-6}
FEN1 (1:1 ratio)		64	1000	1×10^{-3}		100	1000	3.5×10^{-4}
PKM2 (1:1 ratio)		800	1000	1×10^{-3}		512	1000	1×10^{-4}
VDR (1:1 ratio)		300	1000	1×10^{-7}		512	1000	5×10^{-5}
ALDH1 (1:1 ratio)		2048	1000	1×10^{-4}		256	1000	1×10^{-4}
GBA (1:6 ratio)	Fine-tuning with RBF	60	2000	1×10^{-3}	Classification with the single-layer network	90	2000	2.5×10^{-3}
KAT2A (1:6 ratio)		25	2000	5×10^{-4}		10	2000	1×10^{-5}
ESR1 antagonist (1:6 ratio)		23	2000	5×10^{-4}		10	2000	9×10^{-6}
MAPK1 (1:6 ratio)		50	1000	1×10^{-12}		95	2000	1×10^{-9}
FEN1 (1:6 ratio)		128	1000	1×10^{-3}		512	1000	1×10^{-5}
PKM2 (1:6 ratio)		256	1000	1×10^{-4}		1024	1000	1×10^{-5}
VDR (1:6 ratio)		1024	1000	1×10^{-9}		1024	1000	1×10^{-5}
ALDH1 (1:6 ratio)		2048	1000	1×10^{-7}		512	1000	1×10^{-5}

TABLE B6. Applied hyperparameters for the quantum-pretrained embedding condition 1 (with the classical neural network from the XYZ feature map NQE)

Dataset (class label)	Embedding step	Batch size	Iteration	Learning rate	Embedding step	Batch size	Iteration	Learning rate
GBA (1:1 ratio)	Fine-tuning with RBF	25	1000	1×10^{-5}	Classification with the single-layer network	45	1000	5×10^{-6}
KAT2A (1:1 ratio)		32	2000	1×10^{-5}		40	1000	8×10^{-4}
ESR1 antagonist (1:1 ratio)		64	1000	1×10^{-4}		25	1000	1×10^{-6}
MAPK1 (1:1 ratio)		50	2000	1×10^{-7}		100	1000	1×10^{-5}
FEN1 (1:1 ratio)		50	2000	1×10^{-5}		100	1000	1×10^{-6}
PKM2 (1:1 ratio)		32	2000	1×10^{-5}		90	1000	1×10^{-7}
VDR (1:1 ratio)		64	1000	2.3×10^{-3}		1024	1000	1×10^{-4}
ALDH1 (1:1 ratio)		256	1000	1×10^{-3}		2048	1000	5×10^{-4}
GBA (1:6 ratio)	Fine-tuning with RBF	26	1000	7.5×10^{-3}	Classification with the single-layer network	32	1000	6×10^{-3}
KAT2A (1:6 ratio)		30	1000	1×10^{-3}		512	1000	9×10^{-5}
ESR1 antagonist (1:6 ratio)		64	1000	1×10^{-3}		200	1000	7×10^{-6}
MAPK1 (1:6 ratio)		64	1000	1×10^{-4}		512	1000	8×10^{-6}
FEN1 (1:6 ratio)		25	1000	1×10^{-3}		2048	1000	1×10^{-3}
PKM2 (1:6 ratio)		32	1000	1×10^{-3}		100	1000	1×10^{-5}
VDR (1:6 ratio)		256	1000	1×10^{-3}		512	1000	5×10^{-5}
ALDH1 (1:6 ratio)		256	1000	1×10^{-3}		1024	1000	5×10^{-5}

TABLE B7. Applied hyperparameters for the quantum-pretrained embedding condition 2 and 3 (with the classical neural network from the ZZ feature map NQE)

Dataset (class label)	quantum-pretrained embedding condition	Batch size	Iteration	Learning rate	quantum-pretrained embedding condition	Batch size	Iteration	Learning rate
GBA (1:1 ratio)	quantum-pretrained 2	95	1000	1×10^{-10}	quantum-pretrained 3	110	2000	1×10^{-7}
KAT2A (1:1 ratio)		150	2000	1×10^{-12}		80	2000	1×10^{-4}
ESR1 antagonist (1:1 ratio)		50	2000	5×10^{-9}		90	2000	2×10^{-4}
MAPK1 (1:1 ratio)		30	2000	1×10^{-9}		80	2000	7×10^{-4}
FEN1 (1:1 ratio)		100	1000	5×10^{-9}		256	1000	1×10^{-4}
PKM2 (1:1 ratio)		256	1000	1×10^{-8}		128	1000	1×10^{-5}
VDR (1:1 ratio)		128	1000	1×10^{-5}		512	1000	1×10^{-4}
ALDH1 (1:1 ratio)		512	1000	1×10^{-4}		1024	1000	1×10^{-5}
GBA (1:6 ratio)	quantum-pretrained 2	50	2000	1×10^{-10}	quantum-pretrained 3	100	2000	1×10^{-3}
KAT2A (1:6 ratio)		90	2000	1×10^{-12}		80	2000	1.4×10^{-4}
ESR1 antagonist (1:6 ratio)		130	2000	7×10^{-10}		90	2000	5×10^{-4}
MAPK1 (1:6 ratio)		90	2000	1×10^{-9}		50	2000	9×10^{-5}
FEN1 (1:6 ratio)		150	1000	2×10^{-10}		64	1000	1×10^{-6}
PKM2 (1:6 ratio)		500	1000	5×10^{-7}		512	1000	1×10^{-5}
VDR (1:6 ratio)		512	1000	1×10^{-5}		1024	1000	1×10^{-4}
ALDH1 (1:6 ratio)		1024	1000	1×10^{-4}		2048	1000	1×10^{-4}

TABLE B8. Applied hyperparameters for the quantum-pretrained embedding condition 2 and 3 (with the classical neural network from the XYZ feature map NQE)

Dataset (class label)	quantum-pretrained embedding condition	Batch size	Iteration	Learning rate	quantum-pretrained embedding condition	Batch size	Iteration	Learning rate
GBA (1:1 ratio)	quantum-pretrained 2	64	1000	1×10^{-3}	quantum-pretrained 3	25	1000	1×10^{-4}
KAT2A (1:1 ratio)		32	1000	9×10^{-4}		128	1000	5×10^{-3}
ESR1 antagonist (1:1 ratio)		16	1000	5×10^{-5}		13	1000	5×10^{-5}
MAPK1 (1:1 ratio)		110	1000	5×10^{-6}		128	1000	5×10^{-7}
FEN1 (1:1 ratio)		100	1000	1×10^{-5}		128	1000	1×10^{-6}
PKM2 (1:1 ratio)		100	1000	1×10^{-8}		256	1000	2×10^{-7}
VDR (1:1 ratio)		300	1000	5×10^{-6}		512	1000	1×10^{-10}
ALDH1 (1:1 ratio)		2048	1000	1×10^{-7}		2048	1000	1×10^{-8}
GBA (1:6 ratio)	quantum-pretrained 2	50	2000	1×10^{-10}	quantum-pretrained 3	64	1000	5×10^{-4}
KAT2A (1:6 ratio)		90	2000	1×10^{-12}		200	1000	2×10^{-3}
ESR1 antagonist (1:6 ratio)		130	2000	7×10^{-10}		550	1000	2.1×10^{-3}
MAPK1 (1:6 ratio)		90	2000	1×10^{-9}		256	1000	1×10^{-4}
FEN1 (1:6 ratio)		150	1000	2×10^{-10}		512	1000	1×10^{-4}
PKM2 (1:6 ratio)		500	1000	5×10^{-7}		2048	1000	1×10^{-4}
VDR (1:6 ratio)		512	1000	1×10^{-5}		256	1000	9×10^{-5}
ALDH1 (1:6 ratio)		1024	1000	1×10^{-4}		1024	1000	1×10^{-3}

TABLE B9. The classical neural network architectures applied for the quantum (NQE with the ZZ feature map) and classical embedding model

Dataset (class ratio)	Architecture (Linear layers and activation functions)
GBA (1:1 ratio)	Linear(39, 4096)/ReLU/Linear(4096, 2048)/ReLU/Linear(2048, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
KAT2A (1:1 ratio)	Linear(39, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
ESR1 antagonist (1:1 ratio)	Linear(39, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
MAPK1 (1:1 ratio)	Linear(39, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
FEN1 (1:1 ratio)	Linear(39, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
PKM2 (1:1 ratio)	Linear(39, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
VDR (1:1 ratio)	Linear(39, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
ALDH1 (1:1 ratio)	Linear(39, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
GBA (1:6 ratio)	Linear(39, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
KAT2A (1:6 ratio)	Linear(39, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
ESR1 antagonist (1:6 ratio)	Linear(39, 2048)/ReLU/Linear(2048, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
MAPK1 (1:6 ratio)	Linear(39, 2048)/ReLU/Linear(2048, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
FEN1 (1:6 ratio)	Linear(39, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
PKM2 (1:6 ratio)	Linear(39, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
VDR (1:6 ratio)	Linear(39, 2048)/ReLU/Linear(2048, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)
ALDH1 (1:6 ratio)	Linear(39, 1024)/ReLU/Linear(1024, 512)/ReLU/Linear(512, 256)/ReLU/Linear(256, 128)/ReLU/Linear(128, 64)/ReLU/Linear(64, 32)/ReLU/Linear(32, 16)

TABLE B10. The classical neural network architectures applied for the quantum (NQE with the XYZ feature map) and classical embedding model

Dataset (class ratio)	Architecture (Linear layers and activation functions)
GBA (1:1 ratio)	Linear(39, 30)/ReLU/Linear(30, 15)
KAT2A (1:1 ratio)	Linear(39, 240)/ReLU/Linear(240, 120)/ReLU/Linear(120, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)
ESR1 antagonist (1:1 ratio)	Linear(39, 120)/ReLU/Linear(120, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)
MAPK1 (1:1 ratio)	Linear(39, 30)/ReLU/Linear(30, 15)
FEN1 (1:1 ratio)	Linear(39, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)
PKM2 (1:1 ratio)	Linear(39, 30)/ReLU/Linear(30, 15)
VDR (1:1 ratio)	Linear(39, 30)/ReLU/Linear(30, 15)
ALDH1 (1:1 ratio)	Linear(39, 240)/ReLU/Linear(240, 120)/ReLU/Linear(120, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)
GBA (1:6 ratio)	Linear(39, 30)/ReLU/Linear(30, 15)
KAT2A (1:6 ratio)	Linear(39, 120)/ReLU/Linear(120, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)
ESR1 antagonist (1:6 ratio)	Linear(39, 240)/ReLU/Linear(240, 120)/ReLU/Linear(120, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)
MAPK1 (1:6 ratio)	Linear(39, 120)/ReLU/Linear(120, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)
FEN1 (1:6 ratio)	Linear(39, 480)/ReLU/Linear(480, 240)/ReLU/Linear(240, 120)/ReLU/Linear(120, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)
PKM2 (1:6 ratio)	Linear(39, 240)/ReLU/Linear(240, 120)/ReLU/Linear(120, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)
VDR (1:6 ratio)	Linear(39, 240)/ReLU/Linear(240, 120)/ReLU/Linear(120, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)
ALDH1 (1:6 ratio)	Linear(39, 120)/ReLU/Linear(120, 60)/ReLU/Linear(60, 30)/ReLU/Linear(30, 15)