DRU Model for Universal Quantum Classifier

Sadi Nicolas Mendoza Soriano

August 2025

Input:

Let the data matrix be

$$\mathbb{X} = \{\mathbf{x}_n \in \mathbb{R}^P : n \in \{1, \ldots, N\}\}, \quad \mathbb{X} \in \mathbb{R}^{N \times P},$$

where N denotes the number of samples (rows) and P the number of features (columns). The n-th sample is denoted by $\mathbf{x}_n \in \mathbb{R}^P$, corresponding to the n-th row of $\mathbb{X}$.

Each sample $\mathbf{x}_n$ is associated with a label $\mathbf{y}_n$, so that the label set is defined as

$$\mathbf{Y} = \{\mathbf{y}_n \in \mathcal{C} : n = 1, \ldots, N\}, \quad \mathbf{Y} \subseteq \mathcal{C}^N,$$

where $\mathcal{C} = \{c_1, \ldots, c_C\}$ is the set of possible classes and $C = |\mathcal{C}|$ represents the total number of classes.

—

Encoding:

The encoding strategy in the model depends on the dimension of each sample $\mathbf{x}_n \in \mathbb{R}^P$, and is divided into two cases:

(i) If $P < 3$, we set $P' = 3$ and introduce the transformation

$$\mathcal{P} : \mathbb{R}^P \longrightarrow \mathbb{R}^{P'},$$

such that the set of transformed vectors is given by

$$\mathcal{X} = \left\{\mathcal{P}(\mathbf{x}_n) : n = 1, \ldots, N\right\} \in \mathbb{R}^{N \times P'}.$$

Explicitly, the transformation is defined as

$$\mathcal{P}(\mathbf{x}_n) = \begin{bmatrix}\mathbf{x}_n & \Theta_{3-P}\end{bmatrix}, \quad \Theta_{3-P} = (0, \ldots, 0) \in \mathbb{R}^{3-P}.$$

where Θ_{3-P} is a zero-vector that completes $\mathbf{x}_n$ to reach dimension $P^{'}$.

(ii) If $P > 3$, we define

$$K = \left\lceil \frac{P}{P'} \right\rceil, \quad P' = 3,$$

and each sample $\mathbf{x}_n$ is encoded as a set of sub-vectors given by the transformation

$$\mathcal{S} : \mathbb{R}^P \to \mathbb{R}^{K \times P'}$$

where K is the number of sub-vector from the input vector $\mathbf{x}_n$ such that the set of transformed vectors is given by

$$\mathcal{X}' = \left\{\mathcal{S}(\mathbf{x}_n) : n = 1, \ldots, N\right\} \in \mathbb{R}^{N \times K \times P'}.$$

then we have that the transformation is

$$\mathcal{S}(\mathbf{x}_n) = \left\{\mathbf{x}_{n,(j)} \in \mathbb{R}^{P'} : j = 1, \ldots, K\right\}, \quad \mathcal{S}(\mathbf{x}_n) \in \mathbb{R}^{K \times P^{'}}$$

where each $\mathbf{x}_{n,(j)}$ corresponds to a contiguous block of dimension P' extracted from $\mathbf{x}_n$. If P is not divisible by P', the last sub-vector is padded with zeros to reach the required dimension, analogously to the transformation defined in case (i).

—

Parametric Gates:
To represent the inputs in a parameterized quantum circuit (PQC), we require a parametrization of single-qubit unitary operators. The parameters $\boldsymbol{\phi}_s = (\phi_1, \phi_2, \phi_3) \in \mathbb{R}^3$ correspond to the rotation angles on the Bloch sphere, defined as:

$$RZ(\phi_1) = e^{-i\frac{\phi_1}{2}\sigma_z}, \quad RY(\phi_2) = e^{-i\frac{\phi_2}{2}\sigma_y}, \quad RZ(\phi_3) = e^{-i\frac{\phi_3}{2}\sigma_z},$$

where σ_y and σ_z are the corresponding Pauli matrices:

$$\sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \sigma_y, \sigma_z \in \mathcal{L}(\mathcal{H}) \cong \mathbb{C}^{2\times 2}.$$

According to the Euler decomposition, any single-qubit unitary operator can be written in its reduced form as:

$$U(\phi_1, \phi_2, \phi_3) = RZ(\phi_1)\, RY(\phi_2)\, RZ(\phi_3),$$

which provides parametrization in terms of three real parameters. Consequently, we define the set of parameterized unitary operators as

$$\mathcal{U} = \left\{ U(\boldsymbol{\phi}_s) \in SU(2) \;:\; \boldsymbol{\phi}_s \in \mathbb{R}^3 \right\},$$

where $U(\boldsymbol{\phi}_s) \in \mathbb{C}^{2\times 2}$ is unitary, satisfying $U(\boldsymbol{\phi}_s)^\dagger U(\boldsymbol{\phi}_s) = I, \quad \det\big(U(\boldsymbol{\phi}_s)\big) = 1$, and acts on the single-qubit Hilbert space

$$\mathcal{H} = \mathbb{C}^2.$$

—

DRU Model
Let $\mathbf{x}_n \in \mathbb{R}^P$ be the feature vector corresponding to the n-th sample. After the preprocessing and scaling stage, $\mathbf{x}_n$ is partitioned into subvectors of fixed dimension $P' = 3$, that is,

$$\mathbf{x}_n = \big\{\mathbf{x}_{s,n}\big\}_{s=1}^{K}, \qquad \mathbf{x}_{s,n} \in \mathbb{R}^{P'}, \qquad K = \left\lceil \frac{P}{P'} \right\rceil.$$

If P is not divisible by P', the last subvector is padded with zeros (*zero-padding*) to reach the required dimension.

Let $\mathcal{Q} = \{q_0, \ldots, q_{N_q-1}\}$ denote the set of available qubits, with cardinality $N_q = |\mathcal{Q}|$.

Given the layered structure of the DRU model, each subvector $\mathbf{x}_{s,n}$ modulates a unitary operator on a qubit, defined as

$$U(\boldsymbol{\phi}_{s,n}) \in \mathrm{SU}(2), \qquad \boldsymbol{\phi}_{s,n} \in \mathbb{R}^3,$$

whose parameters are obtained through the affine definition

$$\boldsymbol{\phi}_{s,n} = \mathbf{w}_s \circ \mathbf{x}_{s,n} + \boldsymbol{\theta}_s,$$

where $\mathbf{w}_s, \boldsymbol{\theta}_s \in \mathbb{R}^3$ are trainable parameters, and $\circ$ denotes the Hadamard (elementwise) product.

We define the parameter vector associated with each block (s, n) as

$$\boldsymbol{\phi}_{s,n} = \big(\phi_{s,n}^{(1)}, \phi_{s,n}^{(2)}, \phi_{s,n}^{(3)}\big), \quad \{\phi_{s,n}^{(i)} : i = 1, 2, 3\} \subset \mathbb{R},$$

where each component $\phi_{s,n}^{(i)} \in \mathbb{R}$ is a real scalar representing an angle derived from the feature vector in the qubit Hilbert space $\mathcal{H}$.

According to the Euler decomposition, the parameterized operator is expressed as

$$U(\boldsymbol{\phi}_{s,n}) = RZ\big(\phi_{s,n}^{(1)}\big)\, RY\big(\phi_{s,n}^{(2)}\big)\, RZ\big(\phi_{s,n}^{(3)}\big),$$

where $\boldsymbol{\phi}_{s,n}$ corresponds to the vector of angles associated with block (s, n).

For each sample $\mathbf{x}_n \in \mathbb{R}^P$ and each qubit $q_m \in \mathcal{Q}$, we define an operator

$$L^m(\mathbf{x}_n) = \prod_{s=1}^{K} U(\boldsymbol{\phi}_{s,n}^{m}), \qquad L^m(\mathbf{x}_n) \in \mathcal{U}(\mathcal{H}_m),$$

where the product corresponds to the sequential application (ordered in s) of the operators associated with the K blocks.

Finally, since the DRU algorithm requires N_L layers acting on each qubit q_m, we define the set of operators:

$$\mathbb{L} = \{L_i^m(\mathbf{x}_n) : i = 1, \ldots, N_L, \ \ m = 1, \ldots, N_q\}, \quad \mathbb{L} \subset \mathcal{U}(\mathcal{H}_m).$$

where N_L is the number of layers per qubit and N_q is the number of qubits in the model. Given these definitions, the total model acting on a qubit q_m is defined as

$$\tilde{L}^m = \prod_{i=1}^{N_L} L_i^m, \qquad \tilde{L}^m \in \mathcal{U}(\mathcal{H}_m),$$

while the unitary operator acting on the multi-qubit space $\bigotimes_{m=1}^{N_q} \mathcal{H}_m$ is defined as

$$\mathcal{U} = \bigotimes_{m=1}^{N_q} \prod_{i=1}^{N_L} L_i^m,$$

which transforms a basis state $|0\rangle^{N_q}$ into a final state $|\psi_f\rangle^{N_q}$ as

$$|\psi_f\rangle^{N_q} = \mathcal{U}\,|0\rangle^{N_q}.$$

Explicitly, each layer is written as

$$L_i^m(\mathbf{x}_n) = \prod_{s=1}^{K} U(\boldsymbol{\phi}_{s,n,i}^m),$$

with affine parameters given by

$$\boldsymbol{\phi}_{s,n,i}^m = \mathbf{w}_{s,i}^m \circ \mathbf{x}_{s,n} + \boldsymbol{\theta}_{s,i}^m,$$

where the trainable parameters are grouped into the sets

$$\Omega = \{\mathbf{w}_{s,i}^m : s = 1, \ldots, K, \ i = 1, \ldots, N_L, \ m = 1, \ldots, N_q\}, \quad \mathbf{w}_{s,i}^m \in \mathbb{R}^3,$$

$$\beta = \{\boldsymbol{\theta}_{s,i}^m : s = 1, \ldots, K, \ i = 1, \ldots, N_L, \ m = 1, \ldots, N_q\}, \quad \boldsymbol{\theta}_{s,i}^m \in \mathbb{R}^3.$$