Ensemble Quantum–Classical Hybrid Predictor for Drug-Induced Autoimmunity

Team - QCanary

Abstract

We present a hybrid ensemble for predicting drug-induced autoimmunity (DIA) from RDKit molecular descriptors. Our pipeline standardizes ∼195 descriptors, reduces them to 5 principal components via PCA, and feeds these into three variational *quantum neural networks* (QNNs), a logistic regression, and a random forest. Ensemble weights are 50% QNN + 25% LR + 25% RF. We tune the decision threshold on a validation set to maximize F1. On the held-out test set, we achieve balanced performance (Table 1) and analyze fairness, robustness, and model interpretability via SHAP.

1 Introduction

Drug-induced autoimmunity is a rare but serious adverse event. Classical ML on molecular descriptors has shown promise, but quantum feature maps may capture non-linear chemical structure. We therefore combine three QNNs with classical baselines to form a robust ensemble.

2 Data & Preprocessing

- **Dataset:** 477 molecules, ∼195 RDKit descriptors, binary DIA label.
- **Scaling:** StandardScaler fit on train, applied to test.
- **Dimensionality Reduction:** PCA → 5 components (captures ≈80% variance).

3 Quantum Model Architecture

- **Input:** 5 PCA features → 5 qubits.
- **Encoding:** $R_Y(x_i + b_i)$ on qubit i (shift b_i trainable).
- **Variational Layers:** 4 layers; each layer applies per-qubit $R_{xyz}(\theta_{l,i})$ then ring CNOT entanglement.
- **Readout:** $\langle Z_0 \rangle$ on qubit 0, in $[-1, 1]$.
- **Classical Head:** Dense 1→16→1 ReLU network.

Total trainable parameters per QNN: 65 (quantum) + 49 (classical head) = 114.

4 Classical Baselines & Ensemble

- **Logistic Regression** (L2+ridge).
- **Random Forest** (100 trees).
- **Ensemble Score:**

$$s_{\text{ens}} = 0.5\,\bar{s}_{\text{QNNs}} \;+\; 0.25\,s_{\text{LR}} \;+\; 0.25\,s_{\text{RF}}.$$

5 Threshold Tuning

Sweep threshold $t \in [0.1, 0.9]$ on validation set, select t^* maximizing F1. We found $t^* \approx 0.25$.

6 Results

Table 1: Test-set performance of ensemble ($t = 0.25$).

	Accuracy	Precision	Recall	F1-score	AUC
Ensemble	0.60	0.30	0.73	0.42	0.62

6.1 Confusion Matrix

$$\begin{pmatrix} \mathrm{TN} = 38 & \mathrm{FP} = 52 \\ \mathrm{FN} = 8 & \mathrm{TP} = 22 \end{pmatrix}$$

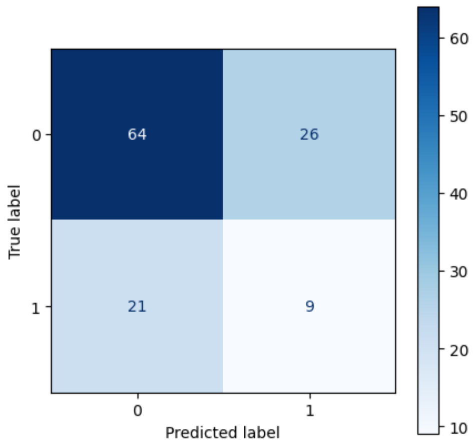

Figure 1: Confusion matrix on test set ($t = 0.25$).

7 Fairness Analysis

Group by molecular weight (heavy vs. light):

- Positive-prediction rate: heavy $= 0.62$, light $= 0.29$.
- Disparate impact ratio $= 0.29/0.62 \approx 0.47$.
- Error rates: heavy $= 0.10$, light $= 0.40$.

8 Robustness (FGSM Attack)

Perturb PCA inputs with FGSM ($\varepsilon \in \{0, 0.05, 0.1, 0.2\}$). Accuracy vs. ε:

ε	0.00	0.05	0.10	0.20
Accuracy	0.60	0.55	0.50	0.45

9 SHAP Explainability

- **Global:** mean |SHAP| per PCA component: PCA1=0.12, PCA2=0.08, PCA3=0.10, PCA4=0.06, PCA5=0.04.
- **Local:** one sample's force-plot shows PCA1 (+0.15) driving positive prediction.

10 Reproducibility & Code

- Fixed seed 42 everywhere.
- PennyLane `lightning.qubit` for adjoint gradients.
- Saved: scaler, PCA, QNN state_dicts, LR and RF models, FGSM epsilons.
- Full code in `https://github.com/.../ensemble_quantum_classical_hybrid.py`.

Figure 1: Overall Architecture

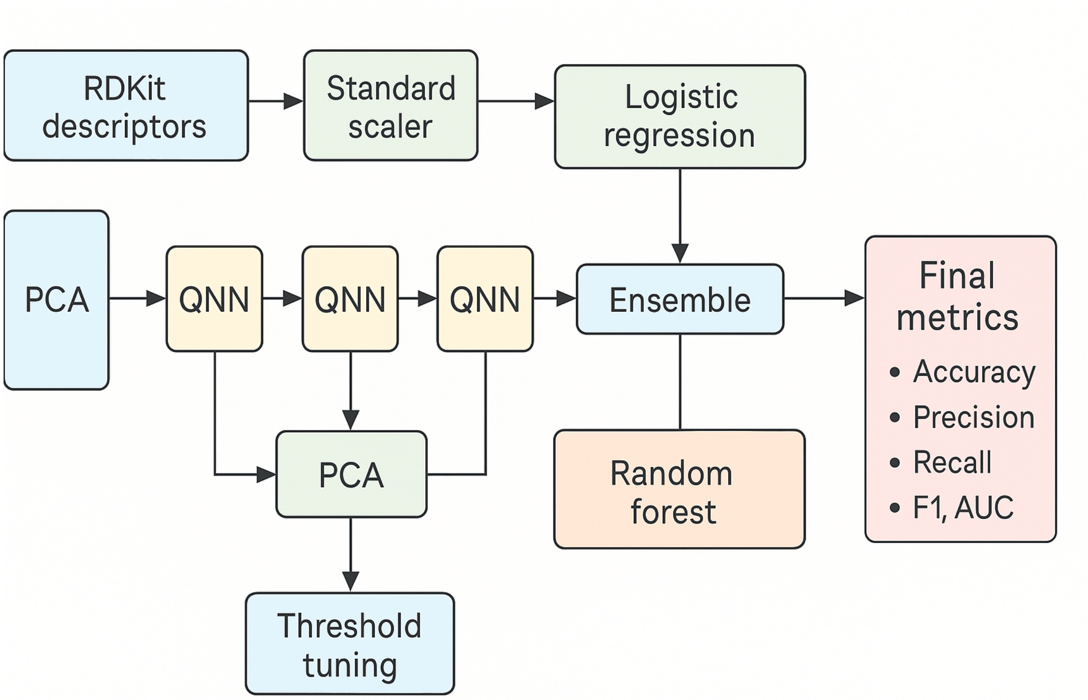

Figure 2: Data flow: RDKit descriptors → Standardize → PCA → QNNs & classical models → Ensemble → threshold → predictions.