Google Colab: Medical Data Benchmark

Public Wisconsin Diagnostic Breast Cancer dataset | November 2, 2026

This summary describes a classical machine-learning benchmark run in Google Colab using a public research dataset. It is not a diagnostic study and does not use patient records provided by DeTection Inc.

The dataset contained 569 samples and 30 numeric features. A fixed, stratified 80/20 split (random_state=42) yielded 455 training and 114 test samples. StandardScaler was fitted on the training set only.

The full 30-feature logistic-regression baseline reported 98.25% test accuracy, F1 score 0.9861 and ROC AUC 0.9954. A separate eight-component PCA plus logistic-regression benchmark reported 95.61% accuracy, F1 score 0.9645, ROC AUC 0.9944; PCA retained 92.84% of variance.

These figures reflect one held-out split. Further replication and appropriate evaluation are required before drawing general conclusions.

Research limitations: These are exploratory public-dataset software simulations. No physical QPU was used in the reported Amazon Braket local-simulator study; no quantum advantage or clinical validation is established. No patent specifications are disclosed.