This report presents a supervised learning analysis of EEG-derived quantum potential and p‑adic entropy features. We built and compared five classifiers—Random Forest, Logistic Regression, Support Vector Machine (SVM), K‑Nearest Neighbors (KNN), and Gradient Boosting—to distinguish between two subject groups (subset\_1 vs. subset\_2) in dataset. Performance was assessed via test accuracy, 5‑fold cross‑validation, confusion matrices, and ROC/AUC curves.

**Data Overview**

* **Samples:** 131
* **Features (4):**
  + quantum\_rest
  + padic\_entropy\_rest
  + quantum\_task
  + padic\_entropy\_task
* **Target:** subset (binary: subset\_1, subset\_2)
* **Identifier:** subject (dropped before modeling)
* **Missing values:** None

A blue rectangular bars with numbers

AI-generated content may be incorrect.The classes are reasonably balanced (about 57% vs. 43%), so no special balancing techniques were applied.

**Preprocessing & Methodology**

1. **Label Encoding:** Converted subset to 0/1.
2. **Train/Test Split:** 80% train, 20% test (27 samples held out).
3. **Scaling:** Features standardized via StandardScaler.
4. **Models Trained:**
   * Random Forest (100 trees)
   * Logistic Regression (L2 penalty)
   * SVM (RBF kernel)
   * KNN (k=5)
   * Gradient Boosting (default scikit‑learn settings)
5. **Evaluation Metrics:**
   * Test accuracy & classification report
   * 5‑fold cross‑validation accuracy
   * Confusion matrix
   * ROC curves & AUC

**Random Forest Baseline**

* **A screenshot of a computer

  AI-generated content may be incorrect.Test Accuracy:** 92.6%
* **5‑Fold CV Accuracy:** 90.1%
* **Precision / Recall / F1:**
  + Class 0: 0.94 / 0.94 / 0.94
  + Class 1: 0.91 / 0.91 / 0.91

**A graph with red squares

AI-generated content may be incorrect.Key Insights:**

* **quantum\_task** is the dominant predictor (≈ 48%).
* **padic\_entropy\_rest** (23%) and **quantum\_rest** (17%) also contribute meaningfully.
* **padic\_entropy\_task** has the smallest, but non‑negligible, importance (10%).

**Multi‑Model Comparison**

**Confusion Matrices**

**A screenshot of a graph

AI-generated content may be incorrect.A screenshot of a graph

AI-generated content may be incorrect.**

**A screenshot of a graph

AI-generated content may be incorrect.A screenshot of a graph

AI-generated content may be incorrect.A screenshot of a graph

AI-generated content may be incorrect.**

* **Logistic Regression** and **Gradient Boosting** both achieved **96.3%** test accuracy, each misclassifying only 1 out of 27 samples.
* **Random Forest** and **KNN** each misclassified 2 samples; **SVM** misclassified 3.

**Cross‑Validation & AUC**

**A graph showing a number of blue rectangular objects

AI-generated content may be incorrect.**

* **AUC Scores:**
  + Random Forest: 0.96
  + Gradient Boosting: 0.95
  + Logistic Regression: 0.94
  + KNN: 0.91
  + SVM: 0.85
* **A graph of a curve

  AI-generated content may be incorrect.**
* **CV Accuracy:**
  + RF & LR: ≈ 90.1%
  + GB & KNN: ≈ 87.0%
  + SVM: ≈ 86.3%

**Conclusions & Recommendations**

1. **Top Models:**
   * **Logistic Regression** offers the best balance of simplicity and performance (96.3% test accuracy, 90.2% CV).
   * **Random Forest** and **Gradient Boosting** are close runners‑up—both highly robust (AUC ≈ 0.95–0.96).
2. **Feature Insights:**
   * **quantum\_task** is the single most informative feature, warranting deeper neurophysiological investigation.
   * The rest/rest vs. task feature split suggests state‑dependent EEG dynamics are key discriminators.