

# Adversarial Robustness of ECG-Lens and Deep Learning Models in ECG Classification

Assignee Research

June 7, 2026

## Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: What is the robustness of ECG-Lens and other deep learning models to adversarial perturbations in ECG signals, and how does this compare to traditional machine learning classifiers in terms of. 17 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.4/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: ECG-Lens: Benchmarking ML & DL Models on PTB-XL Dataset. Research question: What is the robustness of ECG-Lens and other deep learning models to adversarial perturbations in ECG signals, and how does this compare to traditional machine learning classifiers in terms of AUC-ROC on PTB-XL and extended MIT-BIH datasets?.

## 2 Methodology

Systematic literature search across multiple databases yielded 14 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.4/10.

## 3 Results

14 papers retrieved. 17 claims extracted; 1 independently verified. Quality review score: 4.4/10.

## 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.



## 5 Extracted Claims

Claim	Verified	Confidence
Decision Trees can be used to recursively partition the feature space in order to make accurate classifications, and the	×	0.05
The decision tree model performed reasonably well, resulting in a moderate accuracy score of 52%.	×	0.04
The decision tree model has a relatively low F1 score of 22% and ROC-AUC score of 58%.	×	0.08
Logistic Regression is widely used as a linear classification algorithm that models the likelihood of an outcome.	×	0.06
Accuracy is an important metric that determines the overall accuracy of a model's predictions, providing a broad view of	×	0.03
In situations where class imbalance is present, as is commonly seen with ECG data, relying completely on accuracy can be	×	0.05
Supplementary metrics, such as ROC-AUC, F1-score, precision, and recall, are utilized to gain a more detailed understand	×	0.09
The formula for accuracy is $\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$ .	×	0.01
The ROC-AUC can be calculated from the ROC curve, which charts the true positive rate against the false positive rate fo	×	0.03
Precision examines the ratio of true positive instances among positive predictions, whereas recall assesses the model's	×	0.02
The formula for precision is $\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$ .	×	0.01
The ECG-Lens model achieved an accuracy of 80%, ROC-AUC of 90%, F1-Score of 78%, precision of 80%, and recall of 76%.	✓	0.16
The LSTM model achieved an accuracy of 73%, ROC-AUC of 87%, F1-Score of 72%, precision of 78%, and recall of 71%.	×	0.08
The CNN model achieved an accuracy of 71%, ROC-AUC of 85%, F1-Score of 69%, precision of 73%, and recall of 66%.	×	0.08
The Random Forest Classifier achieved an accuracy of 54%, ROC-AUC of 62%, F1-Score of 23%, precision of 25%, and recall	×	0.10
The Logistic Regression model achieved an accuracy of 40%, ROC-AUC of 50%, F1-Score of 21%, precision of 21%, and recall	×	0.09
The PTB-XL dataset contains five classes: NORM (7293 records), MI (4103 records), ST/T Change (3869 records), CD (3790 r	×	0.08

## References

- <http://arxiv.org/abs/2604.15822v1>
- <http://arxiv.org/abs/2008.07651v1>
- <http://arxiv.org/abs/1809.00958v1>