

Ensemble Methods vs. Deterministic Models in High-Dimensional Tabular OOD Detection

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: Do ensemble methods like BE-SNNs outperform single deterministic models in OOD detection on high-dimensional tabular datasets using metrics like AUC-ROC, and how does this compare to other Bayesian. 18 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Contextual Learning for Anomaly Detection in Tabular Data. Research question: Do ensemble methods like BE-SNNs outperform single deterministic models in OOD detection on high-dimensional tabular datasets using metrics like AUC-ROC, and how does this compare to other Bayesian approaches?.

2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.2/10.

3 Results

15 papers retrieved. 18 claims extracted; 0 independently verified. Quality review score: 3.2/10.

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
The contextual learning framework was evaluated on eight datasets: Bank Marketing, Beth Cybersecurity, Census, CMC, KDD,	×	0.09
The datasets vary significantly in size, feature count, and anomaly prevalence.	×	0.03
The Bank Marketing dataset has a Kvcc score of 0.210, a Khet score of 0.343, a Kins score of 0.016, and a Kfnl score of	×	0.02
The Beth Cybersecurity dataset has the highest overall complexity score of 46.264.	×	0.02
The CMC dataset has the lowest overall complexity score of 1.000.	×	0.02
The DSVDD method ranks 7th for the Bank Marketing dataset.	×	0.01
The CWAE method ranks 1st for the Beth Cybersecurity dataset.	×	0.01
The WAE method ranks 9th for the Spotify dataset.	×	0.01
The average rank of the DSVDD method across all datasets is 7.375.	×	0.03
The average rank of the CWAE method across all datasets is 2.75.	×	0.03
The CWAE method achieves the best AUCROC score of 0.695 for the Bank Marketing dataset.	×	0.02
The CWAE Context method achieves the best AUCROC score of 0.997 for the Beth Cybersecurity dataset.	×	0.03
The CWAE method achieves the best AUCROC score of 0.743 for the Census dataset.	×	0.02
The CWAE Context method achieves the best AUCROC score of 0.787 for the CMC dataset.	×	0.02
The CWAE method achieves the best AUCROC score of 0.872 for the KDD dataset.	×	0.02
The CWAE method achieves the best AUCROC score of 1.000 for the LANL dataset.	×	0.02
The CWAE Context method achieves the best AUCROC score of 0.906 for the Solar Flares dataset.	×	0.02
The CWAE method achieves the best AUCROC score of 0.711 for the Spotify dataset.	×	0.02

References

- <http://arxiv.org/abs/2306.11113v2>
- <http://arxiv.org/abs/2507.22692v1>
- <http://arxiv.org/abs/2509.09030v2>