

REVIEWER'S REPORT

Manuscript No.: IJAR-57821

Title: Machine Learning Algorithms in Network Traffic Classification – A Survey

Recommendation:

Accept as it is

Accept after minor revision.....

Accept after major revisionyes.....

Do not accept (*Reasons below*)

Rating	Excel.	Good	Fair	Poor
Originality			Y	
Techn. Quality			Y	
Clarity			Y	
Significance			Y	

Reviewer's ID: JPR- Dr.shaweta sachdeva

Detailed Reviewer's Report

- Tighten the problem statement:** The intro jumps into encryption impact quickly. Add 1 line upfront framing why NTC still matters in 2025-2026 with QUIC, HTTP/3, and IoT surge. Right now it starts with "One of the major developments..." which feels broad.
- KDD dataset context:** You mention using KDD for evaluation. Since KDD is from 1999, reviewers will flag it. Add a comment noting it's for baseline comparison, and that newer datasets like CIC-IDS2017, ISCX-VPN, UNB CIS are covered in the survey tables. This preempts the "outdated dataset" critique.
- Define scope early:** You cover Application ID, Protocol Detection, and Traffic Behavior Analysis. State that upfront so readers don't expect intrusion detection only.
- For Literature Review section**
- Table referencing:** Tables 1, 2, 3 are strong. Add a transition sentence before each: "To compare recent work systematically, Table X summarizes..." Right now tables appear a bit abruptly.
- 2025 papers weighting:** You cite several 2025 works like Kotak et al., Yzzogh et al., Lu et al.. Highlight that as evidence the survey is current. Comment: "This review prioritizes 2023-2025 studies to reflect post-encryption trends."
- Consistency in algorithm names:** You use "Tree-based algorithm" and "C4.5 Decision Trees". Pick one naming style. Suggest: "Decision Tree (C4.5)" for clarity.
- For Research Gaps & Motivation**
- Link gaps to tables:** You list gaps like manual feature engineering, poor generalization, small labeled data. Add a comment tying each gap to a reviewed paper. Ex: "Manual feature engineering remains a bottleneck". Makes the motivation feel grounded.
- Online learning point:** You mention models need to update over time. That's a strong future direction. Suggest adding "concept drift" as the formal term — reviewers like that phrasing.[10][11]
- For Methodology & Results framing**
- NTC pipeline description:** Lines 130-139 describe the NTC process well. Since you don't want headings, use a phrase like "The typical NTC workflow comprises..." to keep flow natural.

REVIEWER'S REPORT

13. **Class imbalance emphasis:** You flag class imbalance as a key issue. Add a comment suggesting future work could test SMOTE, cost-sensitive learning, or focal loss. Shows depth without needing a new subsection.
14. **For Conclusion**
15. **Avoid repetition from abstract:** The conclusion repeats that Random Forest performed best vs. Suggest rephrasing conclusion to focus on implication: "Hence, ensemble methods offer a practical path for deployment in encrypted-traffic environments."
16. **Call to reproducibility:** You mention "reproducible, statistically sound comparison". Strengthen with: "Code and preprocessing details will be released to aid reproducibility." Even if you don't, it signals good practice.
17. **General language/style comments**
18. **Typos to fix:** "studytechniques", "anintegration", "drift duration" should be "flow duration".
19. **Flow vs drift:** You wrote "drift duration". Standard term is "flow duration" or "inter-arrival time." Comment: "Check terminology for consistency with NTC literature."
20. **Citation density:** Intro paragraphs have clusters and. Spread them out or merge to avoid looking like citation stuffing.[1][2][3][4]