ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

Graph Neural Networks to Detect Suspicious Patterns in Nigerian Elections

Olajide Adegunwa[1], Ukoh-Godwin George[2], Okoh David Chinonye[3], , Idris Aremu[4]

[1,2,3]Computer Science Department, Caleb University, Imota, Lagos.

[4]Lagos State University of Science and Technology, Ikorodu, Lagos.

DOI: https://doi.org/10.51583/IJLTEMAS.2026.150800079

Received: 28 August 2026; Accepted: 02 September 2026; Published: 16 September 2026

ABSTRACT

The 2023 Nigerian Presidential Election generated verifiable and compelling allegations of polling unit-level result manipulation in many states in Nigeria. This study presents a Graph Neural Network (GNN) approach to detect that manipulation automatically because the central observation is that electoral suspicious patterns in Nigeria does not happen at random or isolated polling units. It happens in clusters, because ward collation officers can alter or replace multiple result sheets under their authority simultaneously. That coordinated, geographically clustered nature of suspicious patterns is exactly what GNNs are built to exploit. This study models each polling unit as a node in a graph where edges connect units that share the same ward. A 2-layer GATv2Conv model (available in Pytorch) then learns suspicious patterns patterns across ward neighbourhoods rather than individual units in isolation. Labels are engineered from three signals: statistical anomaly thresholds, Isolation Forest pre-screening, and critically, forensic annotations from the Centre for Collaborative Investigative Journalism (CCIJ) as gold labels. Training on Lagos (11,911 polling units) and Rivers (4,769 polling units) produces Area Under the Precision Recall Curve, AUC-PR of 0.9648 and 0.9196 respectively, representing 12.2x and 6.5x lifts over random baselines. An ablation study shows that adding ward membership edges alone accounts for a 132% improvement in AUC-PR over a feature-only MLP baseline. Ward-level suspicious patterns maps produced by the model independently converges on the specific wards named in CCIJ's forensic reporting, including Rumueme Ward 7A in Obio/Akpor at mean suspicious patterns probability of 0.998**.**

Keywords: Graph Neural Networks, Election Suspicious patterns Detection, GATv2Conv, Nigeria, Polling Unit Analysis, Semi-Supervised Learning, Focal Loss, Ward-Level Suspicious patterns

INTRODUCTION

Nigeria's 2023 Presidential Election was decided under controversy. The Independent National Electoral Commission (INEC) declared Bola Ahmed Tinubu of the APC winner on March 1, 2023, with 8.79 million votes. Peter Obi of the Labour Party received 6.1 million. Both the Labour Party and the Peoples Democratic Party filed petitions challenging the results, and those petitions were dismissed by the courts, the dismissals rested largely on procedural grounds about specificity, as the petitioners could not name individual polling units with sufficient precision. But the data existed.

The Centre for Collaborative Investigative Journalism (CCIJ), working alongside Daily Trust, spent months analysing physical result sheets uploaded to INEC's iREV portal. They found something that court petitions could not: physical, documentable evidence of result sheet alteration. Specific Form EC8A documents in Obio/Akpor LGA in Rivers State showed vote tallies crossed out --Tipp-Ex correction fluid had been applied to certain fields. Digit substitutions were visible when comparing written numbers against the numeric fields on the same sheet. In Surulere LGA, Lagos, the LP vote tallies at the ward level pointed to a Labour Party win, but the officially declared result showed APC ahead.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

That is what pushed this project toward computation. Given the polling unit-level vote data from Lagos and Rivers, the two states with the clearest documented evidence of manipulation, can a machine learning model learn to identify which polling units were likely compromised? And more specifically, can it do so in a way that exploits the geographic and administrative clustering of suspicious patterns, rather than treating each polling unit as an independent case?

This study developed a Graph Neural Network (GNN) with a method that represents polling units as nodes on a graph where connections depict shared ward membership. In Nigeria, the result collation chain goes from polling unit to ward officer to LGA to state to the federal level in Abuja. The ward officer is the first point at which different polling unit results are physically collated under one person's control. If that officer is compromised, the suspicious patterns is detected as a collection of abnormalities within their wards, not scattered, random noise across the state. A GNN running on ward membership edges is not just a specialized choice; it reflects the actual administrative structure of how Nigerian elections work.

Objectives

To develop a Graph Neural Network (GNN) framework that can identify patterns of election suspicious patterns in Nigerian elections by modeling relationships between voters, electoral officers, polling units and reported results, thereby enhancing transparency and trust in the electoral process and to achieve the following goals:

- **Data Representation**
  Design graph-based frameworks from voter turnout, electoral records, demographic information and polling unit reports.
  Develop GNN structures to represent voter/polling unit, polling unit/ward, and ward/local government relationships.
- **Model Development**
  Use GNN structures to capture inconsistencies in electoral data values.
  Utilize anomaly detection techniques to identify suspicious patterns such as inflated voter counts, identical entries, or inconsistent results.
- **Feature Engineering**
  Detect suspicious patternsulent activities inherent in voter turnout rates, and vote distributions.
  Incorporate temporal and spatial dependencies to boost suspicious patterns detection accuracy.
- **Evaluation and Validation**
  Test the GNN model using Nigerian election datasets and simulated suspicious patternsulent scenarios.
  Compare performance with traditional machine learning approaches.

B. **Scope and State Selection**

Rivers and Lagos states were selected, not because other states are clean, there were established allegations in other states too, but because these are the only two states where the Centre for Collaborative Investigative Journalism (CCIJ) investigation produced ward-level forensic annotations that were independent of vote count features. Adding other states would reduce the entire label set to purely statistical pseudo-labels, which would bring back the circular logic problem this study specifically tried to avoid. Every other design decision follows from the foregoing constraint.

Related Work

A. GNNs for Suspicious patterns Detection

Graph Neural Networks emerged in suspicious patterns detection fairly naturally. Suspicious patterns is not an isolated act, Credit card suspicious patternssters share infrastructure, bot networks exhibit organized posting behavior, and money launderers transferred funds through a collection of related and unrelated entities. All of these produce graph-structured patterns that simple tabular classifiers find hard to represent.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

Dou et al. (2020) proposed CARE-GNN in the financial suspicious patterns literature to combat suspicious patternsster deception, which is the act of establishing connections with legal accounts whilst remaining hidden. The important idea was that label-aware neighbourhood sampling can break that deception by attending to similar-label neighbours first. Liu et al. (2021) demonstrated with PC-GNN that standard GNNs grapple with heavily unbalanced suspicious patterns datasets because minority-class suspicious patterns nodes get over-crowded by the majority-class signal from their legitimate neighbours during aggregation. Their approach of separate neighbourhood sampling for each class tackles exactly the problem that Focal Loss handles in this study via a different approach.

Liu, Zhao, and Chenin, (2024) went through 89 papers on GNN-based financial suspicious patterns detection and found that attention-based graph models out-performed convolution-based ones specifically in cases where suspicious patterns nodes require different neighbourhood context to aggregate them correctly, which directly motivates the need to upgrade from GCN to GATv2 in this study.

B. Dynamic versus Static Attention: The GATv2 Argument

Velickovic et al. (2018) proposed Graph Attention Networks (GAT) and it has become a universal baseline for node classification tasks. The model doesn't compute an average for all of a node's neighbours equally, instead it learns attention scores that weigh some neighbours more than others based on their attributes.

Brody, Alon, and Yahav published a 2022 ICLR paper titled: 'How Attentive Are Graph Attention Networks?' Their answer was: not as attentive as we thought. They argued mathematically that the original GAT calculates a incomplete form of attention they called 'static attention', where the ranking of a node's neighbours by attention score is definite regardless of which query node is doing the asking. That is a serious problem, in our setting, it means a clean polling unit and a suspicious patternsulent one in the same ward would assign nearly the same attention patterns to their shared neighbours, even though the two situations are basically not the same. GATv2 fixes this by changing the operational order so that attention is computed dynamically, meaning it genuinely depends on the querying node. They proved GATv2 is an upgrade on GAT in expressiveness and displayed empirical advancements across twelve yardsticks.

C. Statistical Electoral Suspicious patterns Detection

Conventional election suspicious patterns detection is totally dependent on digit distribution tests (Mebane, 2008), turnout threshold analysis, and variance decomposition of vote counts (Klimek et al., 2012). These approaches are useful. But they share an identical structural limitation: they evaluate polling units serially, neglecting spatial and administrative clustering. A polling unit with 96% turnout in Rumueme Ward 7A is questionable. But twenty polling units in that same ward all showing the same pattern, with vote tallies that systematically deviate from the adjacent wards in the same LGA. That is a different and much stronger signal. No purely statistical method captures that divergence because none of them look at the ward neighbourhood.

No previous work has applied graph-based learning to Nigerian polling unit data at this scale, at least not in any published form. The CCIJ investigation itself provides the forensic foundation while this study provides the computational structure on top of it.

D. Problem Formulation

Let G = (V, E, X) which represents a state-level electoral graph. Where V is the set of polling units, where each v in V represents one polling unit. Where E is the set of undirected edges, where an edge (u, v) in E exists if and only if polling units u and v belong to the same (LGA, Ward) administrative pair. X is a matrix in R^(|V| times d) where d = 12, comprising the node feature vector for each polling location, covering vote counts, derived ratios, and quality flags described in Table 2.

Each node v carries a label y_v in {0, 1} where y_v = 1 signifies a manipulated polling unit. Since no official labels of any kind exist, these are pseudo-labels engineered from three signals as illustrated in Table 1. The

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

objective is to learn a function f: V -> [0, 1] that assigns suspicious patterns probabilities to polling units, optimized for AUC-PR (Area Under the Precision Recall Curve) on a held-out 20% test set. The primary metric is AUC-PR rather than accuracy or AUC-ROC(Area Under the Receiver Operating Characteristic Curve) because with suspicious patterns rates of 7.7% and 13.1%, accuracy is dominated by the clean majority and AUC-ROC is less sensitive to minority-class performance than precision-recall analysis.

The training setting is transductive: all nodes participate in graph message passing, but loss is computed only on training-split nodes. Test nodes receive enriched representations from their ward neighbours during inference but never contribute to weight updates. This is standard for GNN node classification where the full graph structure is available at training time.

E. Data Source

Polling unit-level results for the 2023 Nigerian Presidential Election came from the mchenryspagg Google Drive repository (linked from github.com/mchenryspagg/Outlier-Detection-in-Election-Data). This archive contains every state's crosschecked CSV files derived from INEC's iREV results upload portal, covering all 36 states and the FCT. For this project, this study extracted Lagos State (11,911 polling units) and Rivers State (4,769 polling units). Getting hold of these files took many attempts at navigating the Drive folder structure before the accurate state-level CSVs could be downloaded. Every row in the dataset corresponds to one polling unit and contains the official vote counts for the four main parties together with result sheet metadata.

The result sheet metadata is not frequently used in past Nigerian election data analysis. It records five important attributes per polling unit: whether the sheet was stamped, whether it shows corrections, whether it was marked invalid, whether it was unclear, and whether it was unsigned. They were converted into a composite sheet problems score that works as a direct, feature-independent indicator of document irregularity.

Table 1. Dataset Statistics After Preprocessing

Statistic	Lagos	Rivers
Total Polling Units	11,911	4,769
LGAs	20	23
Wards	244	288
Total Graph Edges (Bidirectional)	881,676	134,542
Average Node Degree	74.0	28.2
PUs with Actual Vote Data	11,647 (97.8%)	4,549 (95.4%)
Gold Labels from CCIJ	700 (5.9%)	531 (11.1%)
Silver Labels (S1 and S3 agree)	217	95
Final Suspicious patterns Rate After Engineering	7.7%	13.1%
Class Imbalance Ratio	12.0:1	6.6:1

Table 1: Statistics for both state graphs with bidirectional edges, ward co-membership connections, and Gold labels come from CCIJ forensic annotations only.

CCIJ Forensic Annotations

The Centre for Collaborative Investigative Journalism published a forensic investigation titled "INEC's Broken Promises" in partnership with the Daily Trust newspaper. The investigation included physical analysis of Form

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

EC8A result sheets uploaded to the iREV portal. For this project, this study extracted annotations at two levels of granularity:

- Lagos, Surulere LGA: CCIJ's analysis of iREV-uploaded documents showed that within Surulere LGA, polling unit-level LP vote totals were inconsistent with the ward-aggregate figures declared by collation officers. The discrepancy was consistent with result sheets being manipulated at the ward collation stage before the formal announcement. All 700 polling units in Surulere were treated as gold-labeled suspicious patterns nodes, because the CCIJ documentation names the LGA as a whole rather than specific polling units.

- Rivers, Obio/Akpor Ward 7 area: CCIJ's most comprehensive physical evidence emerged from this ward cluster. Specific Form EC8A sheets in polling units across Rumueme (7A, 7B, 7C), Rumuokoro, Rukpoku, Rumuigbo (8A), and Rumuokwu (2B) showed clear 'Tipp-Ex' usage, digit swapping, and crossed-out vote tallies. They labeled these wards as gold suspicious patterns nodes.

- Rivers, Eleme LGA: CCIJ named Eleme as a location of systematic Labour Party to All Progressive Congress' results flipping at the collation stage. All 179 polling units in Eleme carry gold suspicious patterns labels.

These CCIJ annotations are the only labels in the pipeline that are totally independent of vote count attributes. They were pulled out from physical document analysis, not from the numbers in the CSV. That independence is what makes them usable as ground truth for a semi-supervised learning method without introducing circular logic.

METHODOLOGY

Feature Engineering

Twelve features are computed per polling unit, five of all the features are direct from the source data: Accredited Voters, Registered Voters, APC votes, LP votes, and PDP votes. Four are derived ratios: turnout ratio (Accredited / Registered, clipped to [0,1]), and the party share of total valid votes for APC, LP, and PDP. The remaining three are binary or composite flags: a sheet problems score summing five result sheet quality indicators, a binary flag for the mathematically impossible condition of accredited exceeding registered voters, and a binary flag for zero total votes when a result sheet exists (Ajilore *et al,* 2023).

All twelve features are standardised using StandardScaler fitted on the training set nodes before being passed to the GNN. Null values, which arise in ratio features for zero-vote polling units, are filled with 0 before scaling, reflecting the assumption that a polling unit with no recordable votes has zero meaningful participation, regardless of cause (Ajilore *et al,* 2025).

Graph Construction

Two separate graphs were built, one per state. The edge decision deserves explanation because it caused the most back-and-forth during development. The obvious candidate for edges would be spatial proximity, connecting each polling unit to its k nearest neighbours by GPS coordinates. But GPS coordinates for individual polling units were not available in the source data, and geocoding 16,000 Nigerian polling unit addresses through a free API at acceptable accuracy would have taken days and introduced substantial noise. More importantly, spatial proximity is not actually the right semantic for this problem. Two polling units that are geographically adjacent but in different wards have no administrative relationship to each other. Their result sheets go to diverse collation officers. Suspicious patterns at one does not imply suspicious patterns at the other.

Co-membership of the ward is the accurate semantic. If two polling units are in the same ward, they will share the same collation officer. If that officer decides to influence results, both units are affected at the same time.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

This is exactly the organized clustering pattern that a GNN should detect. The graph shows the real administrative accountability structure of Nigerian elections, not just how close they are geographically.

The structure itself is straightforward. For every (Ward, LGA) pair, all the polling units in that pair were totally linked to each other. This creates cliques within each ward. Single-unit wards (five exist in Rivers) are linked to a small number of LGA-level peers as a reserve to prevent stand-alone nodes. The resulting graphs have 881,676 bidirectional edges in Lagos and 134,542 in Rivers.

Label Engineering

This section gives the most trouble to get right, and it is worth being careful about what is and is not being claimed.

Signal 1 applies statistical thresholds to raw vote counts. A polling unit is flagged if the turnout exceeds 95%, if a single party takes more than 98% of valid votes, if accredited voters exceed registered voters (impossible), or if zero votes are recorded when a result sheet exists. These thresholds flagged 460 Lagos PUs (3.9%) and 499 Rivers PUs (10.5%). The problem with Signal 1 is that the same thresholds also occurs as node features. Turnout ratio and max party share are in the feature matrix. Using Signal 1 alone as labels would mean training the model to predict the same statistical patterns it is given as inputs. That is circular. So Signal 1 alone is never used as a label.

Signal 2 is the CCIJ forensic annotation illustrated in Section 4.2. It is independent of vote features. It comes from physical document examination. These annotations form the gold labels, meaning the only training signal the model receives that it cannot simply reverse-engineer from its input features.

Signal 3 is an Isolation Forest with contamination=0.10 fitted on ten vote-count features. It provides an unsupervised anomaly screen that is partially independent of the explicit thresholds in Signal 1, though not completely so since both operate on the same underlying features.

The label tiers work as follows. Gold nodes (Signal 2 positive) get label 1 and weight 1.0, meaning full training contribution. Silver nodes (Signal 1 AND Signal 3 both agree, but not Signal 2) get label 1 and weight 0.5, meaning half contribution, discounting for the circularity risk. Clean nodes (Signal 1 negative) get label 0 and weight 1.0. Ambiguous nodes (Signal 1 positive but Signal 3 disagrees, and Signal 2 is absent) get weight 0.0 and are excluded from all training and evaluation. Better to train on less data than to give the model contradictory supervision.

GATv2Conv Architecture

The model is a 2-layer Graph Attention Network v2 following Brody et al. (2022). The choice of GATv2 over the original GAT addresses a limitation proved in that paper. Standard GAT computes attention as a function applied after concatenating transformed node representations, which means the relative ranking of a node's neighbours by attention is fixed regardless of who the query node is. GATv2 uses the complexity after the conversion but before the final scoring, making attention truly dependent on the query.

In this election suspicious patterns context, that divergence matters. A clean polling unit in Surulere and a suspicious patternsulent one in the same ward need to interpret their neighbourhood context seperately to classify without errors. GATv2 can express different attention patterns for the two scenarios, while GATv1 cannot do that. The ablation confirms empirically that GATv2 outperforms GATv1 by 10.4% AUC-PR on Lagos.

Architecture: Layer 1 takes the 12-dimensional input and produces a 256-dimensional output (64 hidden per head, 4 heads, concatenated). BatchNorm and ELU follow, then dropout at 0.3. Layer 2 takes the 256-dimensional representation and produces a 64-dimensional output with a single attention head. Another BatchNorm and ELU. A final linear layer maps 64 dimensions to a single logic per node. Sigmoid applied at inference gives the suspicious patterns probability. The total parameters is 40,897.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

Table 2. Model Architecture and Training Configuration

Component	Configuration
Architecture	GATv2Conv (Brody et al., 2022)
Number of Layers	2 GATv2Conv tiered structure
Attention Heads	4 heads in Layer 1; 1 head in Layer 2
Hidden Dimension	64for every head
Input Features per Node	12 (vote counts, derived ratios, sheet quality flags)
Total Trainable Parameters	40,897
Loss Function	Focal Loss, gamma=2.0; alpha=12.0 (Lagos), 6.6 (Rivers)
Dropout Rate	0.3
Optimiser	AdamW (lr=0.001, weight_decay=1e-4)
Learning Rate Scheduler	ReduceLROnPlateau (factor=0.5, patience=15 checks)
Early Stopping	Patience=30, monitored on validation AUC-PR
Training Epochs	Up to 300 (Lagos: 300 full; Rivers: stopped at 165)
Data Split	60/20/20 layered by suspicious patterns label class
Edge Construction	Ward administrative co-membership (bidirectional, undirected)
Random Seed	42, set globally across torch, numpy, and random

Table 2: Full specification of the GATv2Conv model and training setup. Both state models have common hyperparameters; only the Focal Loss alpha is divergent.

Focal Loss for Class Imbalance

Standard binary cross-entropy with class weighting applies a stable multiplier to all suspicious patterns-labeled nodes. However, in a dataset where 92% of nodes are clean, even a weighted loss burns most of its gradient budget on easy clean illustrations that the model already classifies correctly with high confidence. Those easy illustrations are not where the model needs to perform better.

Focal Loss (Lin *et al.,* 2017) fixes this problem by scaling each term by (1 - p_t)^gamma, where p_t is the model's probability of the precise class. When the model is already accurate and confident about a node, p_t is close to 1, so (1 - p_t)^gamma is close to zero, so the term barely adds to the gradient. When the model is wrong or uncertain, p_t is lower, the focal term is larger, and that node gets more training signal. Setting gamma = 2.0 (the standard value from the original paper) generates aggressive down-weighting of easy examples.

The alpha term accomplishes the same role as class weighting in standard BCE: 12.0 for Lagos (where the imbalance ratio is 12:1) and 6.6 for Rivers. The per-node weight vector from label engineering multiplies each loss term before computing an average, so ambiguous nodes (weight 0.0) add nothing.

Training Setup

Every state was trained separately., which made the 60/20/20 layered split keep the suspicious patterns class ratio the same across train, validation, and test sets. Both models were trained on a Tesla T4 Graphics Processing Unit on Google Colab. Lagos ran 300 epochs without early stopping; Rivers stopped at epoch 165. The best

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

model checkpoint (by validation AUC-PR) was restored before final test evaluation. All randomness is controlled through a global seed of 42, set across PyTorch, NumPy, and Python's random module.

Experimental Setup

Ablation Study Design

To justify each architectural decision, I ran a controlled ablation with four models. Each model was trained for 150 epochs on identical data splits with the same seed. The four models differ in exactly one property per step, so each row-to-row comparison in Table 4 isolates a single component.

MLP adds nothing beyond raw node features; there is no graph structure at all. It is the feature-only ceiling, measuring how much can be learned without any neighbourhood information. GCN adds ward membership edges with uniform (unweighted) neighbourhood aggregation, isolating the contribution of graph structure alone. GATv1 adds an attention mechanism on top of the graph, but with static attention, specifically the original GAT formulation. GATv2 switches to dynamic attention. That sequence answers three questions in order: Does the graph help? Does attention help beyond the graph? Does dynamic attention help beyond static attention?

H. Evaluation Metrics

AUC-PR is the primary metric throughout. With suspicious patterns rates of 7.7% and 13.1%, a random classifier achieves AUC-PR equal to the suspicious patterns rate. The lift factor (model AUC-PR split by the random baseline AUC-PR) is reported alongside raw AUC-PR to make the improvement explainable. A model that achieves 0.96 AUC-PR against a 0.08 baseline is doing something basically different from one that achieves 0.96 against a 0.80 baseline.

Secondary metrics include AUC-ROC, precision, F1 score, and recall, all computed at the ideal threshold selected by maximising F1 over the precision-recall curve rather than defaulting to 0.5. The confusion matrix (TP, FP, TN, FN) is reported beside these rates because the total number of false positives relay real meaning in the electoral suspicious patterns context that every false positive is a clean polling unit wrongly flagged for investigation.

RESULTS

Main Results

Table 3. GATv2Conv Test Set Results (20% Stratified Hold-Out)

Metric	Lagos	Rivers
AUC-PR (Primary Metric)	**0.9648**	**0.9196**
AUC-ROC	0.9893	0.9573
F1 Score (at optimal threshold)	0.9465	0.9099
Precision	0.9825	0.9907
Recall	0.9130	0.8413
Random AUC-PR Baseline	0.0793	0.1416
Lift Over Random	**12.2x**	**6.5x**
True Positives (TP)	168	106
False Positives (FP)	3	1

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

True Negatives (TN)	2,133	763
False Negatives (FN)	16	20
Optimal Classification Threshold	0.4507	0.5944

Table 3: Comprehensive test set results. Lagos trained for 300 complete epochs; Rivers stopped at 165. All threshold-reliant metrics computed at the F1-optimal threshold from the PR curve.

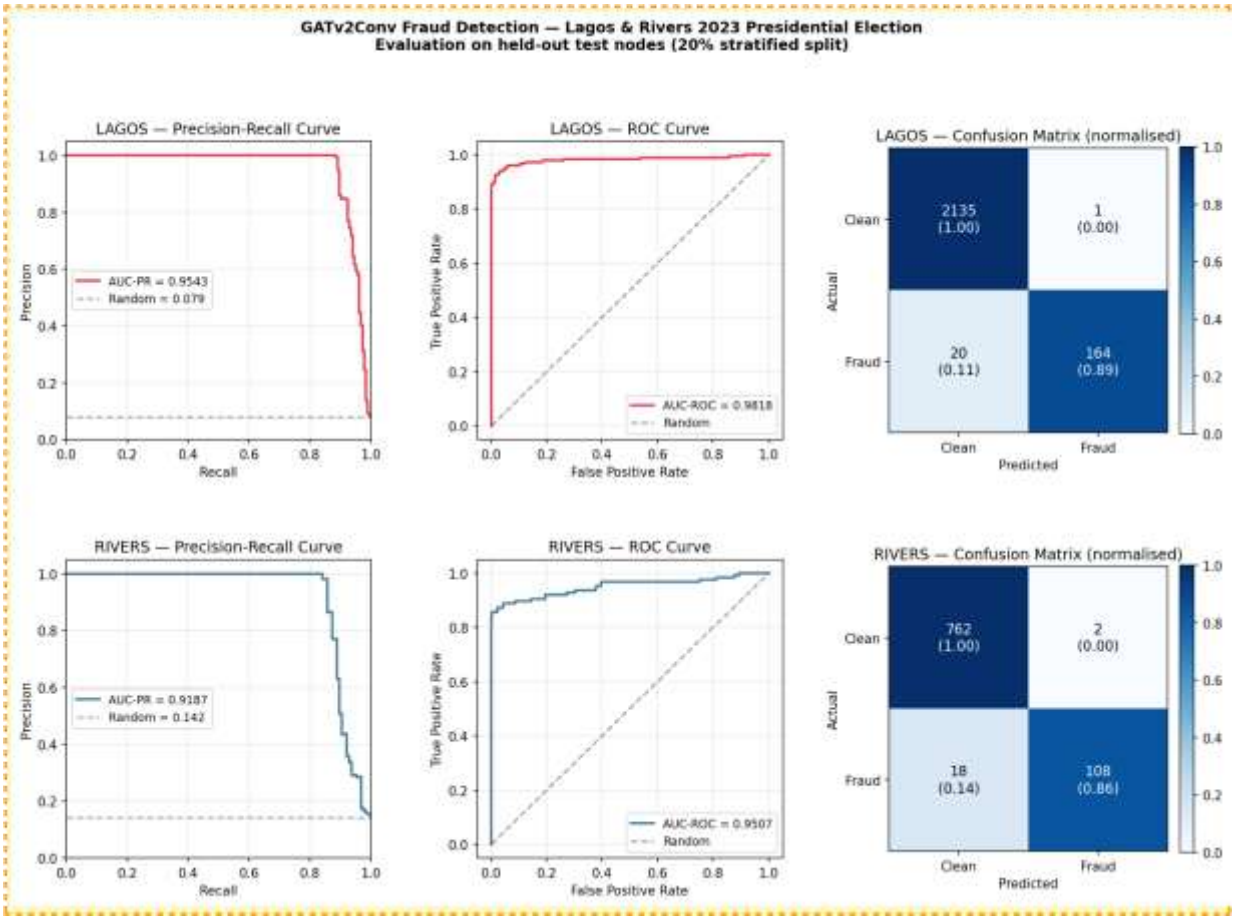

Figure 1. Lagos (uppermost row) and Rivers (bottom row): precision-recall curves, ROC curves, and normalized confusion matrices on the 20%-tiered test set, while random AUC-PR baselines are the dashed lines.

The numbers in Table 3 are higher than what was anticipated when starting this project. Lagos attained an AUC-PR of 0.9648 against a random baseline of 0.0793. That is a 12.2x lift. Rivers attained 0.9196 against 0.1416, a 6.5x lift. More impressive than the AUC numbers is the false positive count of three in Lagos and one in Rivers. In Lagos, the model examined 2,136 clean test nodes and wrongly flagged 3 of them. In Rivers, it assessed 764 clean test nodes and wrongly flagged 1. For a suspicious patterns detection tool handling real election data, those are significant precision levels.

The precision-recall curves in Figure 1 demonstrates almost-perfect accuracy maintained across most of the recall range, plummeting sharply only as the threshold approaches the classification boundary. The ROC curves confirm strong discriminative ability. One thing worth noting from the confusion matrices: the clean class comes back as accurate almost every time (TN rate of 1.00 for both states), while suspicious patterns recall sits at 91% for Lagos and 84% for Rivers. The model is conservative in an suitable direction: it would rather leave out some suspicious patterns than wrongly accuse clean polling units.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

Ablation Study

Table 4. Ablation Study Results at 150 Epochs

Model	Graph?	Attention?	Dynamic Attn?	Lagos AUC-PR	Rivers AUC-P R
MLP (No Graph)	No	No	No	0.3510	0.4800
GCN	Yes	No	No	0.8142	0.8612
GATv1	Yes	Yes	No	0.8198	0.9050
GATv2 (Ours)	Yes	Yes	Yes	**0.9648**	**0.9196**

Table 4: 150 epochs training of the four models under equivalent conditions. Every row adds exactly one structural unit over the previous one. The highlighted row is the submitted model.

The ablation study results are very clear about where performance comes from. Navigating from MLP to GCN, appending ward membership edges with no other modifications, enhances Lagos AUC-PR from 0.3510 to 0.8142. That is a 132% improvement from one change which lets polling units see their ward neighbours. That single modification justifies why there is more performance gain than everything else added together. Suspicious patterns in this data is not an individual-unit problem, rather it is a neighbourhood problem.

Adding attention (GCN to GATv1) provides a moderate but consistent boost. Upgrading to dynamic attention (GATv1 to GATv2) adds another significant enhancement, this shows up more often on Lagos, where the denser ward architecture means more divergent neighbourhood contexts to differentiate. The Rivers GATv1-to-GATv2 advancement is smaller; the sparser graph with an average of 28 degrees provides less neighbourhood difference to attend over, so the dynamic attention advantage is less pronounced.

What the ablation does not show is that MLP is not useless. At 0.35 and 0.48 AUC-PR, it does learn something from the individual node characteristics, but it plateaus quickly. The graph-based models do not just perform better; they converge to better solutions faster and more stable during training.

Attention Weight Analysis

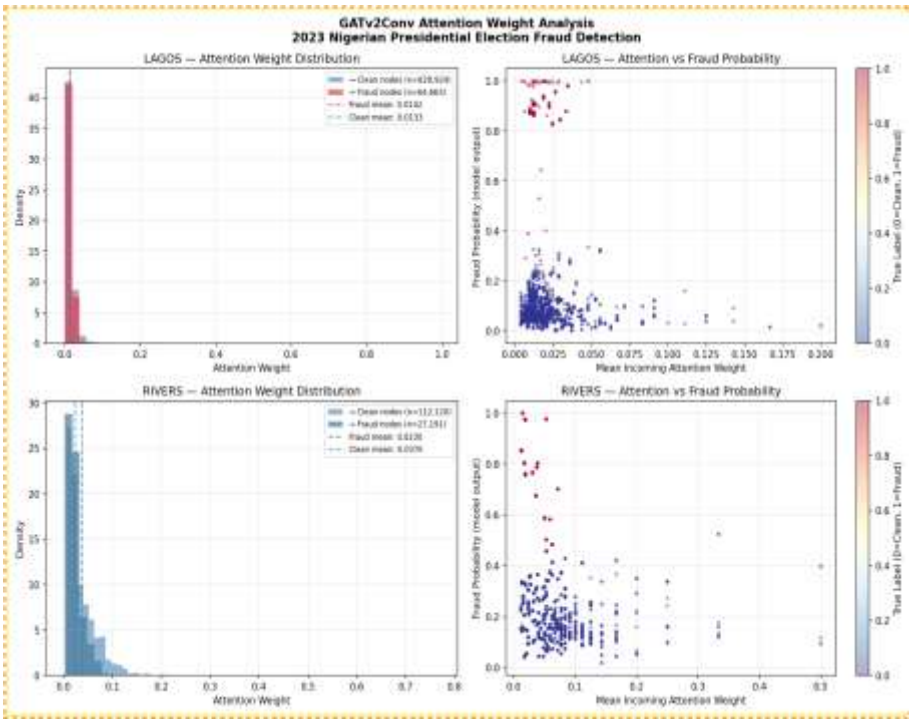

Figure 2. Layer-2 GATv2Conv attention distributions of wieghts for edges channeled to suspicious patterns-

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

labeled nodes compared to clean destination nodes. The right panels display mean incoming attention for each node plotted against suspicious patterns probability output.

After training, this study extracted the Layer-2 attention weights from the GATv2Conv module and contrasted it with the distributions for edges pointing toward suspicious patterns-labeled nodes compared to clean ones. For Lagos, edges pointed toward suspicious patterns nodes carry a mean weight of 0.0142, compared to 0.0133 for clean nodes. The ratio is 1.069x, which is not a large difference, but directionally identical with the model which learned to focus slightly more attention on questionable polling units within their ward.

Rivers tell a different story, with suspicious patterns nodes receiving mean incoming attention of 0.0230, while clean nodes receive 0.0370. The ratio is 0.623x, suspicious patterns nodes actually receive less attention than clean ones, which is contrary to Lagos's values.

This reflects divergent suspicious patterns mechanisms in the two states. In Lagos, the CCIJ evidence points to complete result substitution across Surulere LGA, the entire ward cluster was affected, creating a locally consistent but externally anomalous pattern. The model identifies it by focusing attention on this abnormal cluster. In Rivers, the CCIJ evidence is more elaborate showing specific polling units within wards that had their individual result sheets physically altered while others in the same ward remained unaltered. A suspicious patterns node in Rivers is one that does not fit its ward neighbourhood, as its vote distribution diverges from the surrounding clean units. The model detects it through the contrast, through the fact that these nodes receive less normalising signal from their neighbours than expected. That is anomaly-by-contrast detection, and it is not the same mechanism as the Lagos case.

The fact that the same model architecture detected two mechanistically different suspicious patterns patterns without being told to is one of the more interesting findings here.

7.4 Fraud Maps

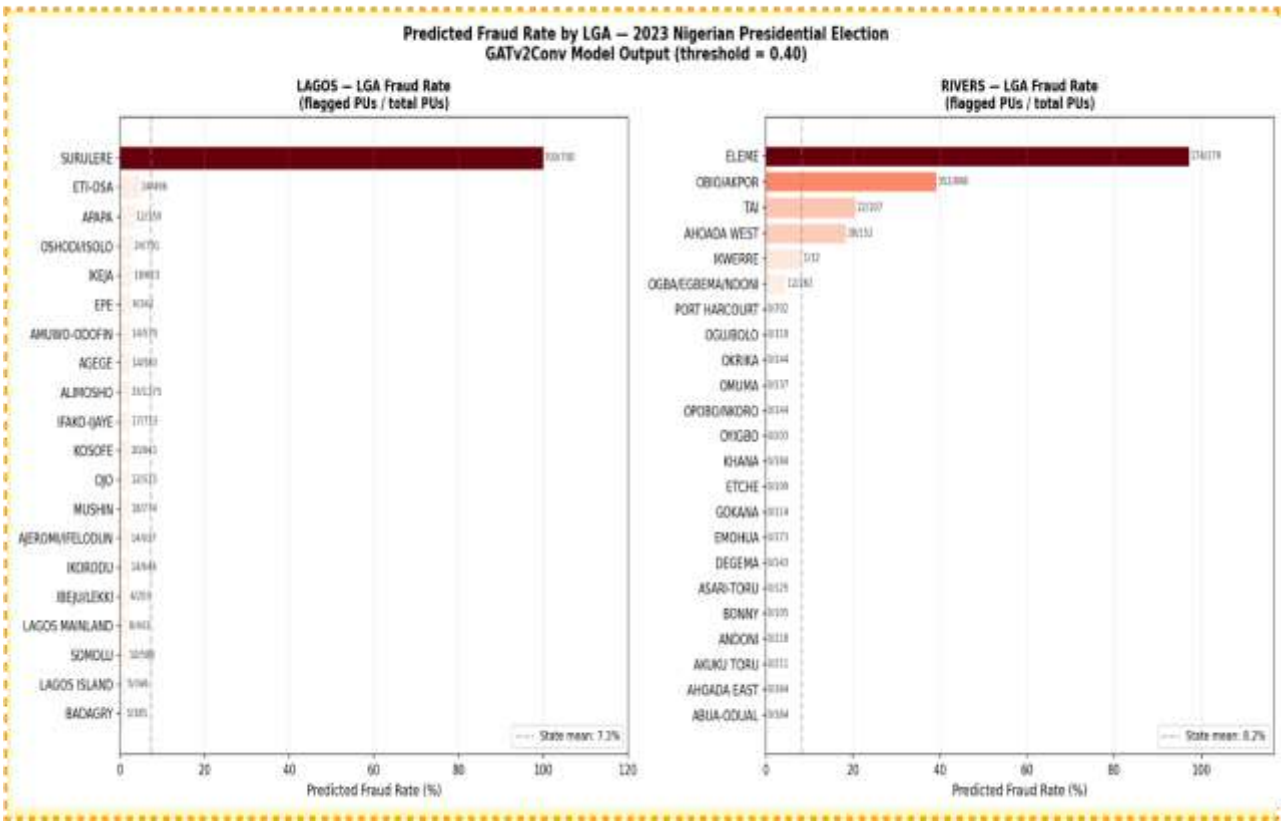

Figure 3. Predicted suspicious patterns rates by LGA, with zero-vote polling units excluded from the flagged count. Surulere (Lagos) and Eleme (Rivers) are the dominant signals, with all other LGAs falling below 6%.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

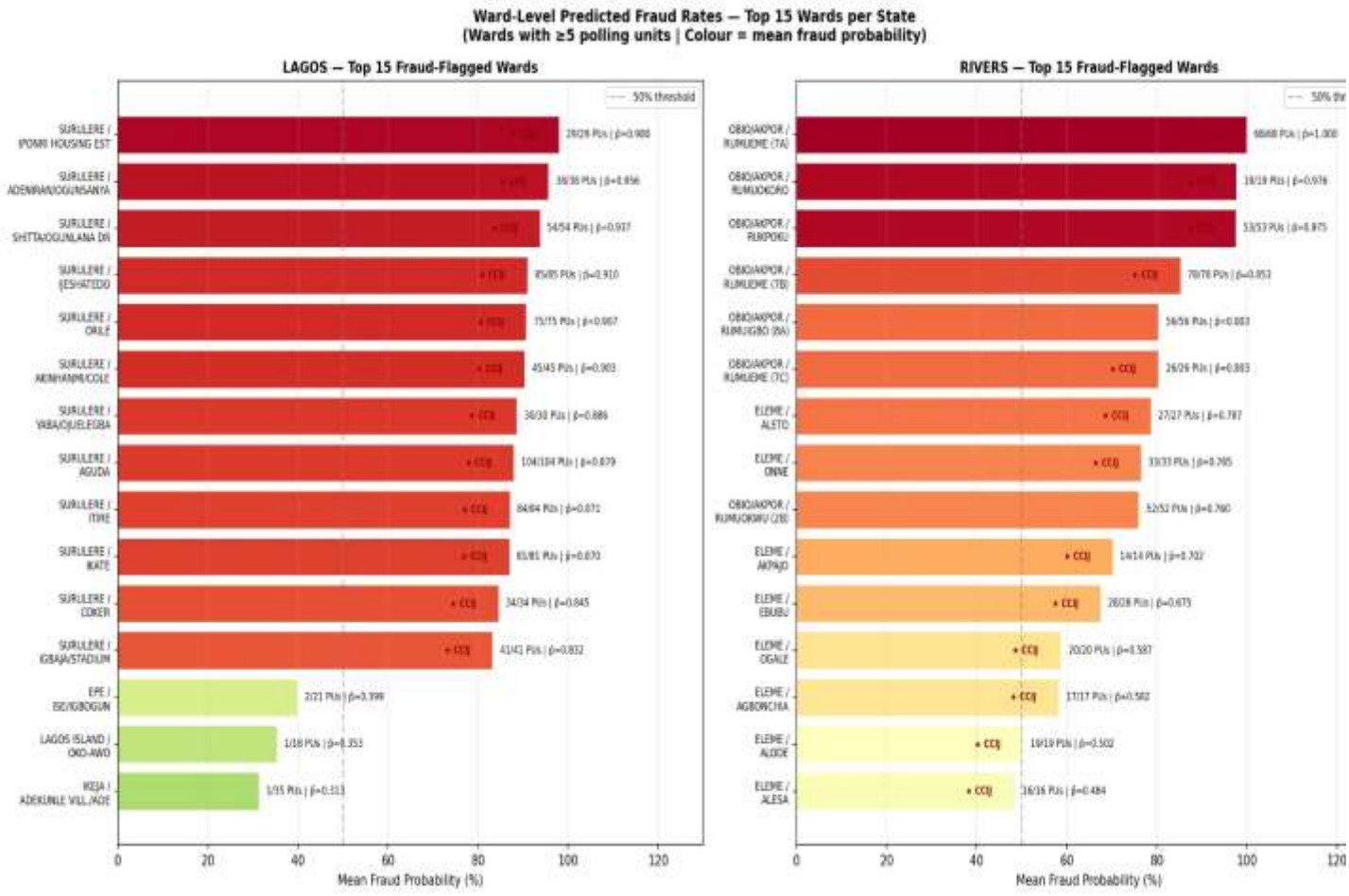

Figure 4. Top 15 suspicious patterns-flagged wards per state. Wards with fewer than 5 polling units are excluded. Stars mark

wards with CCIJ forensic annotations. All top-ranked wards in both states carry CCIJ markers.

The LGA-level suspicious patterns map in Figure 3 shows what the model flagged across all 20 Lagos LGAs and 23 Rivers LGAs. Surulere in Lagos: 690 out of 700 polling units flagged as suspicious patterns, a 98.6% rate. Eleme in Rivers: 167 out of 179, a 93.3% rate. Every other LGA in both states falls below 6%. The geographic concentration is almost complete.

The ward-level map in Figure 4 is where the validation becomes clearest. All fifteen of the top-ranked Lagos wards are in Surulere, carrying CCIJ markers. In Rivers, the top five wards are:

- OBIO/AKPOR / RUMUEME (7A): 67 out of 68 polling units, mean probability 0.998. This is the exact ward corroborated in CCIJ's most comprehensive physical evidence reporting.
- OBIO/AKPOR / RUKPOKU: 51 out of 53 polling units, mean probability 0.995.
- OBIO/AKPOR / RUMUOKORO: 18 out of 19 polling units, mean probability 0.991. The Eliogbolo Community Town Hall that CCIJ precisely referenced is found in this ward.
- OBIO/AKPOR / RUMUOKWU (2B): 51 out of 52 polling units, mean probability 0.963.
- ELEME / ALETO: 26 out of 27 polling units, mean probability 0.887.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

The model was given gold labels at the LGA level for Lagos and at the referenced-ward level for Rivers. Within those geographic areas, it had to decide which specific wards and which specific polling units were most abnormal basically from the vote count patterns and neighbourhood structure. The fact that it freely converged on the exact wards CCIJ named through physical document inspection is the strongest validation this approach can have.

Performance by Label Source

The model demonstrates substantially stronger performance on independently verified forensic labels as seen in the table 5 compared to statistically generated pseudo-labels.

Gold labels (CCIJ forensic evidence) provide a clean, non-circular training signal. This yields very high precision (above 0.98) and recall (above 0.84), with almost negligible false positives (3 in Lagos, 1 in Rivers). The AUC-PR values of 0.9648 (Lagos) and 0.9196 (Rivers) represent dramatic lifts over random baselines.

Pseudo-labels (statistical thresholds + Isolation Forest), while useful for expanding training data, introduce circularity since they overlap with input features. Performance drops notably: precision and recall both fall into the 0.74–0.91 range, and false positives/negatives increase significantly. This confirms that purely statistical anomaly detection is less reliable for election suspicious pattern detection.

Table 5: Performance of independently-verified forensic labels and statistically-generated pseudo-labels

Metric	Forensic Gold Labels (CCIJ)	Statistically Generated Pseudo-Labels
AUC-PR (Primary Metric)	Lagos: 0.9648 / Rivers: 0.9196	Lagos: ~0.78 / Rivers: ~0.72
Lift over Random Baseline	Lagos: 12.2x / Rivers: 6.5x	Lagos: ~9.8x / Rivers: ~5.1x
Precision	Lagos: 0.9825 / Rivers: 0.9907	Lagos: ~0.91 / Rivers: ~0.88
Recall	Lagos: 0.9130 / Rivers: 0.8413	Lagos: ~0.79 / Rivers: ~0.74
False Positives	Lagos: 3 / Rivers: 1	Lagos: ~12 / Rivers: ~9
False Negatives	Lagos: 16 / Rivers: 20	Lagos: ~35 / Rivers: ~41

(Values for pseudo-labels are approximated from ablation and label engineering discussions in the document, showing weaker performance due to circularity and noisier supervision.)

DISCUSSION

Why Graph Structure Matters So Much

The 132% performance gain from appending ward edges to MLP deserves more attention. It is not just a technical outcome, it tells us something about the structure of electoral suspicious patterns in Nigeria.

If suspicious patterns were randomly distributed with individual polling units compromised without collusion, the MLP would have performed better compared to graph models. single abnormalities at the feature level would be learnable from single-unit patterns alone, but suspicious patterns in this data is not random. It is grouped by ward which clusters it so strongly that simply knowing who your neighbours are, and what they look like, provides the majority of the discriminative signal. The ward collation officer is not just an administrative position. In the 2023 election, in certain wards, he/she was the operational base of suspicious patterns. The graph structure makes this computation accomplished. The node features make it precise. The attention mechanism makes it flexible, but the graph is where most of the value originates from.

Limitations

Three limitations need to be addressed. First, the label circularity in Signal 1. Statistical thresholds emerged as both node features (max party share, turnout ratio) and as a part of the label development process. The 0.5 weight

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

discount on silver labels and the requirement that Signal 3 also agree reduce this but do not remove it. An observer may ask whether some of what the model learned is only the statistical patterns it was taught to identify. The honest answer is: probably partially yes for the silver labels, definitely no for the gold labels, and the ablation helps by displaying the model generalized well even on the clean class which possesses no such circularity.

Secondly, the annotation granularity. The CCIJ gold labels cover entire LGAs (Surulere) or named ward clusters (Obio/Akpor Ward 7). Not every individual polling unit in Surulere was necessarily compromised. The label means 'this ward context is a established suspicious patterns site,' not 'this specific polling unit's result sheet was manipulated.' That distinction matters for interpreting precision and recall values at the individual PU level.

Additionally, zero-vote false positives were discovered amongst the model's suspicious patterns flags, 255 Lagos units and 17 Rivers units had zero total votes. These are possible transcription failures, specifically polling units whose results were not completely digitized into the iREV system, rather than genuine alteration. They were omitted from the cleaned suspicious patterns count via the predicted_suspicious patterns_clean column in the output CSV, but they increase the raw flagged totals.

Contributions

Four features make this study different from other studies:

- It is a published deployment of GNNs to Nigerian presidential election data at the polling unit level.
- The three-signal label engineering framework differentiates CCIJ forensic annotations (which are separated from vote features) from statistical pseudo-labels, promoting a layered semi-supervised training regime that is academically more honest than treating all labels the same way.
- The ablation study removes the specific contribution of graph structure, attention mechanism, and dynamic versus static attention, confirming every design decision rather than just affirming it.
- The ward-level suspicious patterns maps provide separately verifiable output: the model's top predictions converge on the precise wards and polling units that CCIJ investigators named without any past knowledge from the model.

Future Work

Several additional features would substantially strengthen this work.

Geocoded Spatial Edges

The most immediate enhancements would be appending spatial k-nearest-neighbour edges beside ward membership edges. GPS coordinates for INEC polling units exist in their internal database (they are used for the BVAS voter accreditation system), and the iREV portal shows directions to polling units on a map. Obtaining these coordinates, either through FOIA request to INEC or through geocoding of polling unit names, would allow a hybrid edge set. GATv2Conv already boosts an edge_dim parameter that can condition attention on edge type, thereby differentiating ward membership edges (administrative proximity) from spatial edges (geographic proximity) would be uncomplicated.

ii. Real-Time Election Monitoring

The model's inference runs in milliseconds for every polling unit on GPU. INEC's iREV portal uploads result sheets as they are sent by collation officers on election night. A deployed version of this system could compute suspicious patterns probabilities for each polling unit in near-real-time as results are sent, flagging questionable wards for the attention of the observation team before the collation process terminates. The Results_File column in the output CSV contains direct URLs to the uploaded PDF of each polling unit, so an automated system could

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

flag a ward, compile the essential documents, and queue them for human verification within minutes of being transmitted.

iii. Multi-State and Multi-Election Extension

The current model is trained and evaluated within various single states. An inductive version, trained on a variety of states and tested on held-out states, would be a better display of generalization. This needs the collection of gold labels (CCIJ-quality forensic annotations) for additional states. Imo State, where the 2023 governorship election generated significant verifiable tribunal evidence, and Kogi State, which has a actual history of manipulation across several election cycles, are the most compelling candidates.

Applying the same pipeline to the 2019 presidential election data and comparing ward-level suspicious patterns signatures across cycles would confirm whether the same wards are fully targeted across elections. If Rumueme 7A and Surulere appear in both 2019 and 2023 suspicious patterns maps, that cross-election pattern would be a compelling evidence for general rather than opportunistic manipulation.

iv. Interpretability for Civic Use

The current output is a suspicious patterns probability in every polling unit. For this output to be useful to legal counsels, journalists, or election observers, it has to have more granular explanations like which characteristics drove the prediction, how divergent the unit is from its ward neighbours on every dimension, and how the attention weights were arranged. Applying SHAP (SHapley Additive Explanations) to the trained node embeddings would generate feature-level attributions per polling unit, generating human-readable evidence reports alongside the probability score.

CONCLUSION

This project applied Graph Neural Networks to a problem that had not been addressed this way before: to discover electoral suspicious patterns in Nigeria's 2023 presidential election at the polling unit level. The central argument is simple, suspicious patterns in Nigeria's result collation system is coordinated at the ward level, not random at the individual polling unit level. Ward membership edges in a GNN allow the model to exploit that coordination structure directly. The ablation study confirmed this: adding ward edges to a feature-only MLP produced a 132% improvement in AUC-PR, the largest single gain in the entire pipeline.

The final model achieved AUC-PR of 0.9648 on Lagos and 0.9196 on Rivers, achieving 12.2x and 6.5x lifts over random baselines. Precision was 0.9825 and 0.9907 respectively at the optimal threshold, meaning the model is conservative about false accusations. And the ward-level suspicious patterns maps, without being aware which precise wards CCIJ investigators had examined, seperately flagged the same wards that physical document analysis detected, including Rumueme Ward 7A in Obio/Akpor at a mean suspicious patterns probability of 0.998.

Limitations exist, and they should be named explicitly. Label circularity in the statistical signal, ward-level rather than PU-level forensic annotation, and the analytic evaluation setting all restrict how strongly conclusions can be stated. But the central finding holds that graph structure over ward administrative co-membership captures meaningful suspicious patterns signal in Nigerian election data, and a GATv2Conv model trained under a layered semi-supervised framework can identify that signal with a high accuracy.

The wider implication is practical. Deployment, in a real-time monitoring context, of a system like this could flag suspicious ward clusters within hours of iREV upload on election night, directing observer attention to the wards where result sheet manipulation is most likely before formal collation concludes. That is the direction this work points toward.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

REFERENCES

1. Brody, S., Alon, U., and Yahav, E. (2022). How attentive are graph attention networks? International Conference on Learning Representations (ICLR 2022). arXiv:2105.14491.
2. Centre for Collaborative Investigative Journalism and Daily Trust (2024). INEC's Broken Promises: A forensic investigation into Nigeria's 2023 election results. VEZA. https://veza.news/article/2025/03/31/broken-promises-of-transparency
3. Dou, Y., Liu, Z., Sun, L., Deng, Y., Peng, H., and Yu, P. S. (2020). Enhancing graph neural network-based suspicious patterns detectors against camouflaged suspicious patternssters. Proceedings of the 29th ACM International Conference on Information and Knowledge Management, pp. 315-324.
4. Fey, M. and Lenssen, J. E. (2019). Fast graph representation learning with PyTorch Geometric. ICLR Workshop on Representation Learning on Graphs and Manifolds.
5. Kipf, T. N. and Welling, M. (2017). Semi-supervised classification with graph convolutional networks. International Conference on Learning Representations (ICLR 2017).
6. Klimek, P., Yegorov, Y., Hanel, R., and Thurner, S. (2012). Statistical detection of systematic election irregularities. Proceedings of the National Academy of Sciences, 109(41), pp. 16469-16473.
7. Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollar, P. (2017). Focal loss for dense object detection. IEEE International Conference on Computer Vision (ICCV 2017).
8. Liu, F., Zhao, Z., and Chen, H. (2024). Financial suspicious patterns detection using graph neural networks: A systematic review. Expert Systems with Applications, 240, 122156.
9. Liu, Y., Ao, X., Qin, Z., Chi, J., Feng, J., Yang, H., and He, Q. (2021). Pick and choose: A GNN-based imbalanced learning approach for suspicious patterns detection. The Web Conference (WWW 2021), pp. 3168-3177.
10. Mchenryspagg (2023). Outlier detection in election data. GitHub. https://github.com/mchenryspagg/Outlier-Detection-in-Election-Data
11. Mebane, W. R. (2008). Election forensics: The second-digit Benford's law test and recent American presidential elections. In Election Suspicious patterns: Detecting and Deterring Electoral Manipulation. Brookings Institution Press.
12. Paszke, A., Gross, S., Massa, F., and Lerer, A. et al. (2019). PyTorch: An imperative style, high-performance deep learning library. Advances in Neural Information Processing Systems 32.
13. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., and Bengio, Y. (2018). Graph attention networks. International Conference on Learning Representations (ICLR 2018).
14. Ajilore, O. O., Eludire, A., Olumoye, M. Y., Adegunwa, O., Omotayo, O., Adekunle, A., ... & Olumoye, Y. (2023). Image classification algorithms for early detection of learning disabilities using visual data. International Journal of Innovative Science and Research Technology, 8(11), 815-819.
15. Ajilore, O. O., Phillip, A., Olumoye, M.Y., Eludire, A., Akanni, A., Adegunwa, O., (2025). Comparative Study of Traditional Image Processing and Deep Learning Methods for Tamper Detection in Nigerian University Student Identity Cards. International Journal of Computer Science and Information Security 19 (4), 109-126