JSCE (Journal of System and Computer Engineering)
Vol.7, No.4, October 2026, pp. 542-554
ISSN (online): 2723-1240
DOI : https://doi.org/10.61628/jsce.v7i4.2972

Research Article Open Access (CC–BY-SA) ■

Coordinate Identity Features for Unsupervised Detection of Attendance Location Fraud

Nur Trismi Auliah [1,a,*]; Wardi [2,b]; Hazriani [3,c]; Syafruddin Syarif [4,d]; Yuyun [5,e]; Imran Taufik [6,f]

[1,3,5,6] Department of Computer Systems, Handayani University of Makassar, Makassar 90222, South Sulawesi, Indonesia
[2,4] Department of Electrical Engineering, Hasanuddin University, Makassar 90245, South Sulawesi, Indonesia
[a] nur.trismi@gmail.com; [b] wardi@unhas.ac.id; [c] hazriani@handayani.ac.id; [d] ssyariftuh376@gmail.com; [e] yuyunwabula@handayani.ac.id; [f] imran_taufik@handayani.ac.id
** Corresponding author*

Abstract

Web-based attendance systems cannot read the mock location flag of a user device, so coordinates produced by spoofing applications pass geofence validation unnoticed. This study asks whether detection can be moved from the device to the historical pattern, and which unsupervised detector suits the data that result. Eleven composite spatio-temporal-behavioural features were engineered from the attendance log of a corporate human resource information system in Indonesia, three of them measuring how often an employee repeats the same coordinate at six-decimal precision. Isolation Forest, a one-class support vector machine and a long short-term memory autoencoder were compared under twelve-fold walk-forward cross-validation on 249,818 records from 212 employees, evaluated against pseudo-labels derived from seven documented attendance rules. The stored models were then applied to 18,711 later records, and the leading indicator was confirmed with ten employees through a blinded matched case-control interview. The one-class support vector machine led on every accuracy metric, reaching a precision-recall area of 0.5367 and a Matthews correlation coefficient of 0.2541, and its ranking held across three calibration scenarios. Removing the coordinate repetition features cost 74.2 percent of the discriminative margin, while the field confirmation separated flagged from comparison employees completely.

Keywords—*attendance fraud, anomaly detection, geofencing, one-class support vector machine, risk scoring*

1. Introduction

Attendance data now drive payroll, productivity reporting and performance review in most large organisations, and smart attendance monitoring has become a research field of its own (Subramanian et al., 2023). Because the falsification of working hours is treated as payroll fraud, and because nobody can read half a million rows by hand and stay consistent, machine learning has been applied to attendance records directly (Zhu et al., 2022). The systems are simple: the server stores time and coordinates, and a geofence, a virtual circle around the office, decides whether the record counts (Prasetyo et al., 2021). Spoofing applications break that logic in one step by feeding the browser a coordinate of the user choosing, and a radius test cannot separate a fabricated point from a genuine one (Arfianto et al., 2024). The gap is architectural: the mock location flag reaches native applications only, which is why the systems able to use it are built as mobile clients (Rajendran et al., 2024) and why countermeasures combining dynamic codes, geofencing and device identifiers assume the same client (Nwabuwe et al., 2023). Satellite positioning, however, has one property that spoofing does not imitate, a metre-level error that

grows in dense urban environments (Han et al., 2024), so two measurements at the same desk never coincide, and at six decimal places an exact repeat by chance sits far below 0.001 percent.

A preliminary examination of 11,858 records from 198 active employees showed why that opening is needed. Late check-ins covered 39.8 percent of the records, incomplete pairs 45.2 percent and short working days 12.1 percent, but the decisive pattern was different: in 35.6 percent of records a single coordinate accounted for more than eighty percent of the attendance of a person, and those records belonged to 62 employees out of 198. Ordinary drift does not produce that, and the four patterns rule out a single-rule solution, since they span location, time and habit at once. Operational logs carry no validated labels, so unsupervised detection is the only honest starting point (Chandola et al., 2009; Ruff et al., 2021), and benchmark evidence shows that which algorithm wins is mostly a question about the data (Han et al., 2022).

Three families of algorithm were therefore tested, because they fail in different ways: Isolation Forest for the tree-based family (Liu et al., 2008), the one-class support vector machine for the kernel-based family (Schölkopf et al., 2001), and the long short-term memory autoencoder for the sequence-based family (Lindemann et al., 2021). Earlier comparisons leave a gap. The forest was faster on network intrusion data while the kernel was more precise in high dimensions (Fuhnwi et al., 2023), a simulation study placed the forest ahead on precision and recall (Agyemang, 2024), and seismic and condition-monitoring work found the two competitive with an efficiency advantage for the forest (Airlangga, 2024; Mattera et al., 2025; Princz et al., 2024), whereas on railway track geometry, where the features are coordinates, the kernel won (Ghiasi et al., 2024). Autoencoders performed well on imbalanced welding data (Zeng et al., 2020) and univariate turbine series (Atsafack et al., 2025), and sequence models have been used against spoofing itself, though with inertial sensors rather than historical logs (Aftatah & Zebbara, 2024; Shafique et al., 2021). None of this work involves human behaviour with a spatial dimension, and the closest study (Rajendran et al., 2024) runs inside a native application, compares no alternative algorithm, and never checks its indicator against outside evidence.

Three things are therefore missing: a comparison of the tree, kernel and sequence families on one location-based attendance log with the training budget and evaluation reference held equal; a feature that treats coordinate repetition at six-decimal precision as a stand-in for the flag a browser cannot read; and a test of the resulting indicator against evidence collected outside the pseudo-labels the researchers themselves wrote. This paper supplies all three: it measures each feature group by ablation, compares the three detectors on accuracy, cost and stability with significance testing and robustness checks, confirms the leading indicator through a blinded matched case-control interview, and recommends one detector as the scoring engine of an attendance dashboard.

Table 1 places this study beside the closest work on five dimensions: data and domain, detectors actually compared, the way location enters the feature set, the evaluation reference, and whether outside evidence was collected. The last row differs on the final three, since no earlier study compares the three paradigms under an equal budget, converts coordinate repetition into a substitute for the inaccessible flag, or tests its indicator against employees blind to their own labels.

Table 1. Position of this study among the most closely related work

Study	Data and domain	Detectors compared	Use of location	Evaluation reference and external check
Zhu et al. (2022)	Attendance records	Supervised classifiers	None	Labelled cases; no field check
Nwabuwe et al. (2023); Arfianto et al. (2024)	Attendance, native client	Rule-based, no detector	Geofence and mock location flag	Rule checks; no field confirmation

Study	Data and domain	Detectors compared	Use of location	Evaluation reference and external check
Rajendran et al. (2024)	Mobile attendance with sensors	One-class support vector machine only	Coordinates read from sensors	Internal rules only
Aftatah and Zebbara (2024)	Navigation signals	Long short-term memory network	Position and velocity	Simulated spoofing scenarios
Ghiasi et al. (2024)	Railway track geometry	One-class support vector machine, Isolation Forest	Geometric coordinates	Recorded track defects
Fuhnwi et al. (2023)	Network intrusion logs	Isolation Forest, one-class support vector machine	None	Benchmark labels
Zeng et al. (2020); Atsafack et al. (2025)	Welding and turbine series	Long short-term memory autoencoder	None	Process labels
This study	Web attendance log, 249,818 records	All three paradigms, equal budget	Coordinate identity at six decimals	Pseudo-labels under three calibrations, plus blinded field confirmation

2. Method

2.1 Problem analysis

The problem has two parts that have to be solved in sequence. The first is representational: a browser cannot observe the authenticity of a coordinate, so authenticity has to be inferred from a pattern that a genuine receiver cannot produce. This study resolves it by moving the unit of observation from the record to the employee history and by engineering features that count how often an identical coordinate recurs. The second is evaluative: an operational log carries no labels, so any claim about accuracy is really a claim about agreement with a reference that the researcher built. This study resolves it in three ways, by deriving every rule indicator from written company regulations rather than from intuition, by re-running the whole procedure under three calibration scenarios, and by confirming the leading indicator with employees who did not know their own labels.

The two solutions are implemented by the workflow shown in Figure 1, which traces the study from the raw attendance log through preprocessing, feature engineering and pseudo-label construction to the three detectors, and from there to evaluation, robustness testing and external validation. The subsections that follow describe each stage in the same order.

2.2 Research design and data

The study is a quantitative comparison run as a computational experiment, with a structured confirmation stage at the end, which makes the overall design an explanatory sequential one. The unit of analysis is dual. Isolation Forest and the one-class support vector machine score records independently, the autoencoder scores a seven-day sequence for each employee, and the confirmation stage works at the level of the person. Three datasets were used. Training used 249,818 records from 212 employees between 15 June 2022 and 31 March 2026, carrying feature engineering, pseudo-label construction, training and evaluation. External validation used 18,711 analysable records from 202 employees between April and July 2026, scored by the stored models with no retraining, of whom 191 passed the data sufficiency check. Confirmation used statements from ten employees, five labelled high risk paired with five labelled low risk and matched on

attendance days and number of records. The window opens on the date the internal attendance rule took effect, so one regulatory regime covers everything analysed.

Three collection techniques were used. Five export files supplied the employee identifier, date, time, attendance type and status, recording source, coordinates at six decimal places, office name, geofence radius and distance to office. Company regulations and the internal memorandum on working hours, attendance recording and the confirmation mechanism supplied the normative basis for every rule indicator. Interviews used a closed instrument written before fieldwork began, and no respondent was told their risk label; blinding here is structural rather than procedural, because the application never displays coordinates to its users.

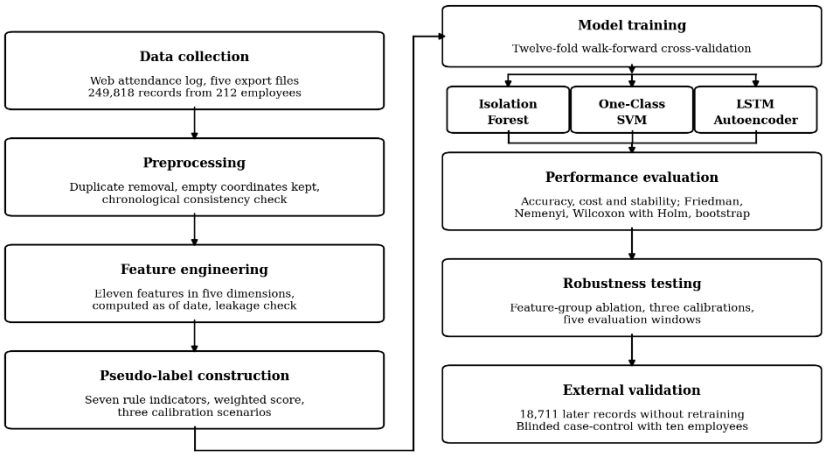


Figure 1. Research workflow from data collection to external validation

2.3 Feature engineering and leakage control

Cleaning removed duplicates keyed on employee, date, time and attendance type, dropped records with no identifier or date, and checked chronological order. Empty coordinates were kept as zero rather than imputed, because a missing coordinate carries information of its own. Eleven features were then derived across five dimensions, as shown in Table 2. The central one is the share of identical coordinates, defined in Equation 1 for employee e up to day t, where C(e,t) is the set of distinct coordinates rounded to six decimal places, the numerator counts the records that fall on the most frequent of those coordinates, and the denominator is the total number of records of that employee up to the same day. Every aggregate feature is computed as of the running date, from history up to that day and nothing later, which is the standard guard against the temporal leakage that inflates reported performance in applied machine learning (Kapoor & Narayanan, 2023). The program re-checks this automatically: cumulative features must be monotonically non-decreasing for each employee, and execution halts if they are not. Scaling used winsorisation, a logarithmic transform and median-interquartile scaling, with every parameter fitted on the training partition of each fold alone.

$$p_{id}(e,t) = \frac{max_{c \in C(e,t)}\, n_c(e,t)}{N(e,t)} \tag{1}$$

Table 2. Composite spatio-temporal-behavioural features derived from the attendance log

Dimension	Feature name	Operational definition
Spatial	Distance from office	Distance from the attendance location to the office reference point in metres, computed with the Haversine formula

Dimension	Feature name	Operational definition
Temporal	Check-in minutes, time deviation, midnight check-in flag, day of week	Check-in time in minutes, its deviation from the standard start of work, an indicator of attendance recorded around midnight, and the day of the week
Behavioural	Daily frequency, no-checkout flag, average working duration	Number of records per employee per day, an indicator of no check-out up to the running date, and the running average of working duration
Coordinate identity	Identical days, identical percentage, maximum streak	Number of days, percentage and longest consecutive run of identical coordinates at six-decimal precision, equal to a resolution of 0.11 metres
Predictive deviation (autoencoder only)	Check-in and location prediction deviation, deviation z-score, consecutive high-deviation days	Absolute deviation of check-in time and attendance distance from the preceding seven-day moving average, normalised against the history of the employee, and the run length of days whose z-score exceeds 2.0

2.4 Pseudo-label construction and calibration

With no anomaly labels in the data, the evaluation reference had to be built. Seven rule indicators were derived from written company provisions, as shown in Table 3. Indicator A3 carries double weight, both because it is the focus of the study and because it is the one indicator with no innocent physical explanation. For a record on day t belonging to employee e, the cumulative rule score and the resulting label are defined in Equation 2, where the indicator function returns one when rule i fires on that record and the weight of each rule is given in Table 3.

$$S(e,t) = \sum_{i=1}^{7} w_i \cdot I_i(e,t) \;,\; y(e,t) = 1 \;\Leftrightarrow\; S(e,t) \geq 2 \tag{2}$$

Thresholds for A1 and A4 were read from percentiles of the training distribution so that they follow the habits of this organisation rather than an outside assumption. Three calibration scenarios were then fixed: a loose scenario at 31.98 percent prevalence, a medium scenario at 22.91 percent and a strict scenario at 10.86 percent. The medium scenario carries the main results, and the other two show how far the conclusions travel.

Table 3. Rule indicators forming the pseudo-labels and their normative basis

Code	Indicator	Weight	Normative basis
A1	Lateness beyond threshold	1	Company regulation on working hours from 08.00 to 17.00
A2	Day without check-out	1	Company regulation on the obligation to record attendance
A3	Identical coordinates (fake positioning)	2	Internal memorandum: attendance buttons activate only inside the office area
A4	Working duration below threshold	1	Company regulation on the eight-hour daily and forty-hour weekly limits
A5	Confirmation source exceeding web source	1	Internal memorandum: attendance confirmation is an exception mechanism
A6	Never recorded a check-out	1	Company regulation on the obligation to record attendance
A7	Never recorded a check-in	1	Company regulation on attendance recording and absence without notice

2.5 Model configuration and training protocol

Isolation Forest used 200 trees, 256 samples at most, and a contamination value equal to the anomalous share of the training fold, bounded to 0.01 and 0.49. The one-class support vector machine used a radial basis kernel, nu equal to the fold contamination, and a gamma chosen per fold from scale, 0.01, 0.001 and 0.0001 by separation on the validation partition. The autoencoder used seven-day sequences, a 32-16-16-32 encoder-decoder, mean squared error with the Adam optimiser, an L2 term of 0.0001 with dropout of 0.2 and recurrent dropout of 0.1, and thirty epochs with early stopping and learning-rate reduction over five seeds per fold. All three were trained under twelve-fold walk-forward cross-validation, which keeps the chronology intact so that no model sees data from after its own test window, and each fold gave every model the same 15,000 records and the same evaluation reference, because otherwise a difference in accuracy could be a difference in budget. Thresholds were set on the validation partition and carried over unchanged, and every run was screened for degenerate prediction.

Raw anomaly scores were mapped to a risk scale from zero to one hundred using calibration bounds stored with the model, as defined in Equation 3, where the raw anomaly score of a record is rescaled between the first and the ninety-ninth percentile of the training score distribution and then clipped to the closed interval. Fixing the bounds at training time, rather than recomputing them for each new batch, is what keeps scores from different assessment periods comparable.

$$R(x) = min\left\{100, max\left\{0, 100 \cdot \frac{s(x) - q_1}{q_{99} - q_1}\right\}\right\} \quad (3)$$

2.6 Evaluation metrics and statistical analysis

Accuracy was read from the F1-score, the precision-recall area with its margin over the prevalence baseline, and the Matthews correlation coefficient, which uses all four cells of the confusion matrix and cannot be inflated by prevalence, and which recent methodological work argues should be preferred to area under the receiver operating characteristic curve for binary problems of this kind (Chicco & Jurman, 2023). Cost was measured as training time per fold and stability as the coefficient of variation of the F1-score, while the Cohen kappa coefficient is reported as an agreement statistic rather than a comparison metric. Significance used the Friedman test at the 0.05 level with the twelve folds as blocks and the Kendall coefficient of concordance as effect size, followed by the Nemenyi post-hoc test drawn as a critical difference diagram, pairwise Wilcoxon signed-rank tests under Holm correction, and bootstrap confidence intervals over ten thousand resamples. Robustness came from three directions, ablating feature groups, re-running the three calibration scenarios, and scoring five evaluation windows of different length. The confirmation stage was read from a two-by-two agreement matrix with Wilson intervals and a two-sided Fisher exact test, which stays valid at low cell frequencies, and model selection followed a decision matrix whose criteria were fixed before any test was run. Everything was implemented in Python with Scikit-learn, TensorFlow-Keras, Pandas, NumPy and SciPy.

3. Results And Discussion

3.1 Spatial characteristics and the limits of geofence validation

Across the 35 office locations, all 205,032 records that carry coordinates fall inside their own geofence, without a single exception. Perfect compliance is in fact the problem, because the only question the geofence asks is whether the transmitted point lies inside the circle. Two employees with almost identical exposure, 422 against 397 attendance days and 426 against 423 records, show what that question misses. The first spread 426 records over 84 distinct coordinates, with 288 of them, 67.6 percent, on one coordinate that repeats to the sixth decimal; the second spread 423 records over 414 coordinates and never put more than two on any of them. With 0.11 metres of resolution and metre-level movement in the signal itself (Han et al., 2024), 288 repeats inside one resolution cell is not something a working receiver produces, and the behaviour is rare rather than general, since the next highest value in the population is 43.7 percent.

3.2 Temporal dependence of the attendance data

Before reading the model comparison, the data were checked for the property a sequence model needs, which is dependence between consecutive observations and not merely an ordering in time. Mean autocorrelation at lag one is 0.361 for check-in time, 0.337 for working duration and 0.238 for distance to office, and the Ljung-Box test rejects the null of no autocorrelation for 88.7 to 95.7 percent of employees. Those numbers look encouraging until the decay is examined. Genuine short-range dependence decays geometrically, so lag seven should be close to lag one raised to the seventh power, that is 0.0008 for check-in time; the observed value is 0.247, roughly three hundred times larger, and for distance to office the curve barely moves, from 0.238 down to 0.180. A flat curve means the structure is not day-to-day dependence but the persistence of individual habit. Two consequences follow, and the second explains most of the results reported below. A gating mechanism has no between-day pattern to learn here, so the sequence family gains nothing structurally. Worse, an autoencoder trained on a persistent series learns to reproduce persistence, and the anomaly of interest, the same coordinate for weeks, is itself persistent, so the model reconstructs it accurately and reports no error.

3.3 Comparative performance and statistical significance

The twelve-fold results are shown in Table 4. The one-class support vector machine takes every accuracy metric, Isolation Forest is second and the autoencoder last, while efficiency runs the other way and not by a small margin: Isolation Forest trains in 1.1 seconds per fold, 4.1 times faster than the kernel-based detector and 224 times faster than the autoencoder. Stability is comparable across all three, between 3.12 and 7.29 percent, none of the 84 runs collapsed into a single class, and against the F1-score of 0.4200 that a degenerate classifier could reach at this prevalence, all three are doing real work.

Table 4. Performance of the three models across twelve folds, reported as mean and standard deviation

Metric	Isolation Forest	One-class support vector machine	Long short-term memory autoencoder
F1-score	0.4612 ± 0.0157	0.4777 ± 0.0149	0.4266 ± 0.0311
Precision-recall area under the curve	0.4899 ± 0.0261	0.5367 ± 0.0485	0.3025 ± 0.0275
Margin over the prevalence baseline	0.2236 ± 0.0463	0.2704 ± 0.0705	0.0362 ± 0.0115
Matthews correlation coefficient	0.1973 ± 0.0607	0.2541 ± 0.0779	0.0853 ± 0.0235
Training time (seconds per fold)	1.1 ± 0.1	4.5 ± 0.5	246.2 ± 18.0
Coefficient of variation of the F1-score	3.40%	3.12%	7.29%
Cohen kappa coefficient	0.1514 ± 0.0780	0.2438 ± 0.0845	0.0457 ± 0.0132

The Friedman test returns p-values between 0.000006 and 0.000431 across the metrics, with Kendall concordance from 0.6458 to 1.0000. A value of 1.0000 on the precision-recall metric and on training time means the three models kept the same order in all twelve folds, without a single inversion. The post-hoc results are shown in Table 5, and the two procedures disagree in an instructive way. Both confirm that the boundary-based detectors beat the autoencoder on everything. For the pair of boundary-based detectors, the Nemenyi test finds a difference only on the precision-recall metric, while the Wilcoxon test under Holm correction finds it on precision-recall and on the Matthews coefficient, which is what one expects from a conservative post-hoc procedure. The rank-biserial correlation of minus 1.000 on the precision-recall metric says the

kernel-based detector was ahead in every fold. All bootstrap intervals exclude zero, the F1-score included, even though neither post-hoc test flagged it, so this is not an artefact of one convenient partition. The same ordering is shown as average ranks in Figure 2.

Table 5. Statistical test results and bootstrap confidence intervals for the Isolation Forest and one-class support vector machine pair

Metric	Friedman p	Kendall coefficient	Nemenyi p	Holm p	Mean difference (95% interval)
F1-score	0.000431	0.6458	0.8134	0.0640	−0.0165 (−0.0297; −0.0031)
Precision-recall area	0.000006	1.0000	0.0380*	0.0015*	−0.0467 (−0.0643; −0.0292)
Matthews correlation coefficient	0.000033	0.8611	0.2317	0.0034*	−0.0569 (−0.0794; −0.0295)
Training time	0.000006	1.0000	–	–	–

An asterisk denotes significance at the 0.05 level. A negative difference indicates a higher value for the one-class support vector machine. The critical difference of the Nemenyi test is 0.957 for twelve blocks and three treatments.

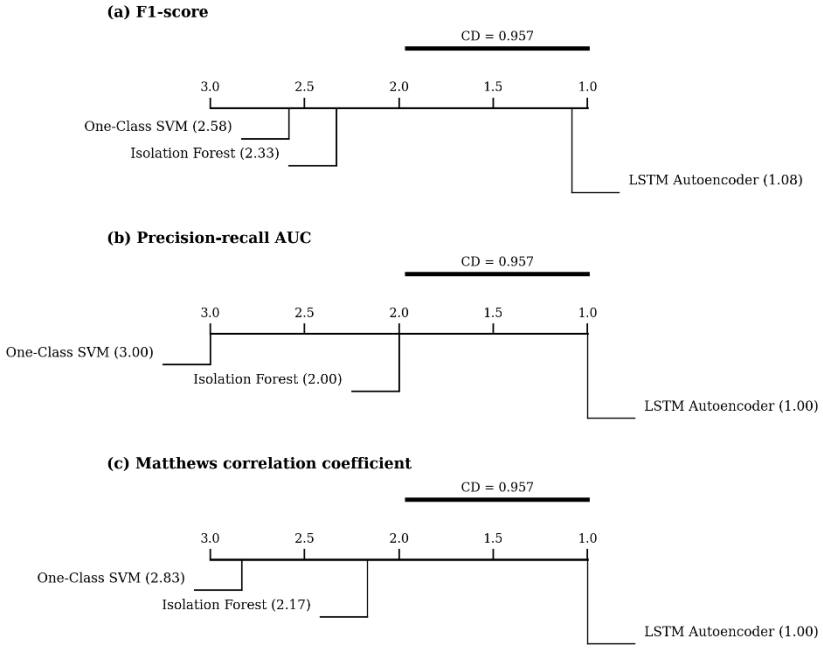


Figure 2. Critical difference diagram of the average ranks of the three models over twelve folds on (a) the F1-score, (b) the precision-recall area under the curve and (c) the Matthews correlation coefficient

3.4 Feature contribution and robustness

Ablation turns the claim about features into a number, as shown in Table 6. The coordinate-identity group carries the detection: removing the three features drops the margin of Isolation Forest over the prevalence baseline from 0.2236 to 0.0577, a loss of 74.2 percent and by far the largest of any variant. The reverse experiment points the same way, since a model built from those three features alone keeps 89.3 percent of the discriminative margin and 97.8 percent of the Matthews coefficient of the full kernel-based model. The other eight features therefore reorder the priority list without deciding who appears on it, which suits them to different jobs, quick screening on one side and ranking on the other. One result was unexpected, dropping the temporal

features raised the precision-recall value of Isolation Forest by 0.0471, because the forest picks split features at random, so a weakly informative feature buys nothing and wastes partitions. They were kept anyway, since they behave differently under a kernel that does no random splitting. Calibration sensitivity told the same story as the folds, the kernel-based detector leads on every metric in all three scenarios, with the Matthews gap running from 0.0477 to 0.1598, so the ranking survives a prevalence that moves from 10.86 to 31.98 percent. That is the strongest statement available here, because without ground truth a study cannot show that its answer is right, only that the answer does not move when the decisions of the researcher move.

Table 6. Results of the feature-group ablation analysis

Feature-set variant	Features	Precision-recall area	Margin over baseline	Matthews coefficient	Change in margin
Isolation Forest, complete	11	0.4899	0.2236	0.1973	reference
Without coordinate identity	8	0.3240	0.0577	0.1899	−0.1659
Without temporal	7	0.5370	0.2707	0.2220	+0.0471
Without behavioural	8	0.4877	0.2214	0.1794	−0.0022
Without spatial	10	0.4874	0.2211	0.1964	−0.0025
Coordinate identity only	3	0.4918	0.2255	0.2169	+0.0019
Kernel-based detector, complete	11	0.5367	0.2704	0.2541	reference
Kernel-based detector, coordinate identity only	3	0.5077	0.2414	0.2484	−0.0290

3.5 External validation and respondent confirmation

The stored models were applied to 18,711 records from April to July 2026 with no retraining, and raw scores were mapped with Equation 3 so that scores from different periods remain comparable. How far each model separates records by severity is shown in Table 7. The kernel-based detector climbs from 8.74 on records with no indicator to 52.50 on records with four, a spread of 43.76 points, with its high-risk share rising from 1.08 to 34.55 percent, and Isolation Forest behaves similarly over a spread of 38.88 points. The autoencoder ends lower at the top of the severity scale than at the bottom, a spread of minus 3.88 points, which is another way of saying it cannot tell a clean record from one that trips four indicators at once.

Table 7. Mean model score by rule score level on the April to July 2026 data

Rule score	Records	Isolation Forest score	Kernel-based detector score	High-risk share	Autoencoder score
0	6,227	18.67	8.74	1.08%	14.30
1	5,843	17.64	10.53	2.62%	15.68
2	5,800	22.11	12.67	3.03%	15.26
3	595	47.63	38.60	24.71%	16.05
4	246	57.55	52.50	34.55%	10.42
Separation range	18,711	38.88	43.76	–	−3.88

At employee level the kernel-based detector flags 8 of the 191 assessed employees, 4.2 percent, two of whom Isolation Forest also flags, while the autoencoder flags nobody; eight names is a caseload a human capital team can work through. Window length turned out to be decisive: monthly partitions give only 16.3 to 20.7 attendance days per person while the coordinate-identity indicator needs at least twenty observation days, and the rank correlation between rule score and model score reaches 0.375 on the combined four-month window against 1.8 to 3.2 times less on the monthly ones. Of the eighteen names produced by the four monthly partitions, only two

survive all four months, so a quarterly window is not a preference but a condition for the list to mean anything.

Ten employees took part in the confirmation stage, paired case-control, none told their label. All five flagged respondents named one fixed spot as the place they record attendance; all five comparison respondents said their spot moves around the office. That is a clean split, with a Cohen kappa of 1.00 and a two-sided Fisher exact p-value of 0.0079, sensitivity and specificity at one hundred percent, and Wilson intervals from 56.6 to 100 percent. The evidence sits outside the rule system, since the event counts come from the database while the statements come from people describing their own habits, and blinding held: three flagged respondents said plainly that they had no idea whether their stored coordinates matched. An application artefact such as rounding or a wireless access point reported as a position would affect everyone in the same office on the same application, and no comparison respondent showed it. The mechanical evidence was the most persuasive part. Of the five respondents who had tried to record attendance from outside the office, the two comparison respondents were refused, which is the area restriction doing its job, while the records of the three flagged respondents, those with 52, 55 and 77 coordinate-identity events, were accepted and stored. On the same system and in the same area, that is only possible if the coordinates reaching the server were not the coordinates of the person sending them.

3.6 Discussion

The ordering of the two boundary-based detectors matches Ghiasi et al. (2024) on railway track geometry, a domain that shares nothing with attendance except the shape of the features, which is the point: on coordinate-bearing data the kernel keeps winning. A one-class support vector machine draws a closed boundary around the region of normal data, so a high share of identical coordinates together with a long streak registers as one pattern, whereas Isolation Forest splits features at random, which suits an extreme value on a single feature and degrades as the signal moves into a combination; ablation shows the same from the other side, since uninformative features hurt the forest and leave the kernel alone. On efficiency the results agree with Fuhnwi et al. (2023), Agyemang (2024) and Mattera et al. (2025), but the practical conclusion differs: training runs periodically, so 3.4 seconds per fold is invisible, while the accuracy gap is not, because on identical data and thresholds the kernel flagged eight employees and the forest two. The forest was kept as a second opinion, since it costs nothing and its agreement with the selected model is a useful consensus signal.

The autoencoder result contradicted expectation: the most sophisticated model finished last on every accuracy metric and showed no separation on new data, and this was not a training failure, since none of the sixty runs degenerated and a 7.29 percent coefficient of variation is unremarkable. The taxonomy of Chandola et al. (2009) explains it better than any tuning argument. Point anomalies dominate here and collective anomalies come next, a coordinate repeated for weeks meaning nothing until read as a run, while contextual anomalies, a person breaking from their own history, are almost absent. Sequence models are built for that third type, so the results of Zeng et al. (2020) and Atsafack et al. (2025) stand but do not transfer to data whose sequential dependence is as weak as subsection 3.2 showed, which demonstrates the argument of Ruff et al. (2021) that model assumptions matter more than architectural depth.

Against Rajendran et al. (2024) three differences stand out, and they are the three empty columns of Table 1: detection moves from the device to the pattern, which is what lets it work on the web and, as the ablation shows, costs little; the kernel advantage is measured against two rival paradigms under an equal budget instead of assumed; and the indicator is tested against evidence collected outside the rules that produced it, a step attendance fraud mitigation has so far left out (Nwabuwe et al., 2023). The limits deserve equal plainness. The best agreement reached is a Matthews coefficient of 0.2541 and a kappa of 0.2438, both weak and both measuring agreement with pseudo-labels rather than fraud, which is the honest basis for treating the output as a

clarification queue rather than a finding. Only one indicator has independent validation, and eight of the ten respondents believed that one record a day is enough to be counted present, so several indicators may measure confusion about the rules rather than misconduct. Because respondents came from the two ends of the score distribution, the perfect sensitivity and specificity describe separation between extremes rather than accuracy across the workforce (Ransohoff & Feinstein, 1978).

4. Conclusions

The composite feature set represents location attendance anomalies well, and the work is done by three of the eleven features, since removing the coordinate-identity group costs 74.2 percent of the discriminative margin, which establishes that coordinate repetition at six-decimal precision can stand in for the mock location flag that a browser cannot read. Among the three detectors the one-class support vector machine leads on every accuracy metric, with a precision-recall area of 0.5367, a margin of 0.2704 over the prevalence baseline and a Matthews correlation coefficient of 0.2541, against 0.4899, 0.2236 and 0.1973 for Isolation Forest and 0.3025, 0.0362 and 0.0853 for the long short-term memory autoencoder; the advantage is statistically significant and holds across twelve folds, three calibration scenarios and five evaluation windows, while the cost ordering runs the opposite way at 1.1, 4.5 and 246.2 seconds per fold, and the autoencoder fails because attendance data are ordered in time without being sequentially dependent. The strengths of the approach are that it needs no sensor access, that its leading indicator was confirmed against evidence from outside the rule system at a p-value of 0.0079, and that its output is a short and actionable list of eight employees out of 191. Its weaknesses are equally clear: agreement with the pseudo-labels remains weak, only one of the seven indicators has independent validation, the confirmation sample came from the two ends of the score distribution and so overstates separation in the middle, the thresholds come from one organisation, and the indicator needs at least twenty observation days, which rules out monthly scoring. The recommendation that follows is to deploy the kernel-based detector as a quarterly risk scoring engine on a zero to one hundred scale, to retain the data sufficiency safeguard, and to publish its output as a clarification queue requiring human verification, while further work extends the confirmation into the middle of the score distribution, tests the thresholds in other organisations, and adds a signal independent of coordinates, such as access point identity, device fingerprints or network addresses.

Acknowledgements

The authors thank the Postgraduate Program of Handayani University of Makassar for its support throughout this study, and the management and human capital team of the company studied, both for access to the attendance log and for allowing their employees to take part in the confirmation interviews. The study received no external research funding.

References

Aftatah, M., & Zebbara, K. (2024). Detecting GPS spoofing attacks using corrected low-cost INS data with an LSTM network. International Journal of Advanced Computer Science and Applications, 15(11), 478–487.

Agyemang, E. F. (2024). Anomaly detection using unsupervised machine learning algorithms: A simulation study. Scientific African, 26, e02386. https://doi.org/10.1016/j.sciaf.2024.e02386

Airlangga, G. (2024). Advanced machine learning techniques for seismic anomaly detection in Indonesia: A comparative study of LOF, isolation forest, and one-class SVM. Jurnal Lebesgue, 5(1), 1–12. https://doi.org/10.46306/lb.v5i1.490

Arfianto, I., Ashari, I. A., & Purwono. (2024). Implementation of geolocator for location manipulation detection in GPS-based attendance application. In Proceedings of the International Conference on Health and Biological Sciences (pp. 124–140).

Atsafack, B. M., Kabiri, C., & Rushingabigwi, G. (2025). A multi-model predictive framework for unsupervised anomaly detection in univariate time series data from hydraulic turbine units. Energy Reports, 14, 4701–4709. https://doi.org/10.1016/j.egyr.2025.11.079

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 1–58. https://doi.org/10.1145/1541880.1541882

Chicco, D., & Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. BioData Mining, 16(1), 4. https://doi.org/10.1186/s13040-023-00322-4

Fuhnwi, G. S., Adedoyin, V., & Agbaje, J. O. (2023). An empirical Internet protocol network intrusion detection using isolation forest and one-class support vector machine. International Journal of Advanced Computer Science and Applications, 14(8), 1–9. https://doi.org/10.14569/IJACSA.2023.0140801

Ghiasi, R., Khan, M. A., Sorrentino, D., Diaine, C., & Malekjafarian, A. (2024). An unsupervised anomaly detection framework for onboard monitoring of railway track geometrical defects using one-class support vector machine. Engineering Applications of Artificial Intelligence, 133, 108167. https://doi.org/10.1016/j.engappai.2024.108167

Han, S., Hu, X., Huang, H., Jiang, M., & Zhao, Y. (2022). ADBench: Anomaly detection benchmark. In Advances in Neural Information Processing Systems (Vol. 35, pp. 32142–32159). Curran Associates.

Han, Z., Wang, X., Zhang, J., Xin, S., Huang, Q., & Shen, S. (2024). An improved velocity-aided method for smartphone single-frequency code positioning in real-world driving scenarios. Remote Sensing, 16(21), 3988. https://doi.org/10.3390/rs16213988

Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. Patterns, 4(9), 100804. https://doi.org/10.1016/j.patter.2023.100804

Lindemann, B., Maschler, B., Sahlab, N., & Weyrich, M. (2021). A survey on anomaly detection for technical systems using LSTM networks. Computers in Industry, 131, 103465. https://doi.org/10.1016/j.compind.2021.103465

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In Proceedings of the 8th IEEE International Conference on Data Mining (pp. 413–422). IEEE. https://doi.org/10.1109/ICDM.2008.17

Mattera, G., Mattera, R., Vespoli, S., & Salatiello, E. (2025). Anomaly detection in manufacturing systems with temporal networks and unsupervised machine learning. Computers and Industrial Engineering, 203, 111023. https://doi.org/10.1016/j.cie.2025.111023

Nwabuwe, A., Sanghera, B., Alade, T., & Olajide, F. (2023). Fraud mitigation in attendance monitoring systems using dynamic QR code, geofencing and IMEI technologies. International Journal of Advanced Computer Science and Applications, 14(4), 938–945. https://doi.org/10.14569/IJACSA.2023.01404104

Prasetyo, W. E., Zaen, A., & Listyorini, T. (2021). Attendance system using QR code and GPS-based geofencing. Journal of Physics: Conference Series, 2111, 012001. https://doi.org/10.1088/1742-6596/2111/1/012001

Princz, G., Shaloo, M., & Erol, S. (2024). Anomaly detection in binary time series data: An unsupervised machine learning approach for condition monitoring. Procedia Computer Science, 232, 1065–1078. https://doi.org/10.1016/j.procs.2024.01.105

Rajendran, R., Vasanth, A., & Kuppusamy, P. (2024). Fake GPS detection in mobile attendance systems using machine learning. Sensors, 24(10), 3154. https://doi.org/10.3390/s24101523

Ransohoff, D. F., & Feinstein, A. R. (1978). Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. New England Journal of Medicine, 299(17), 926–930. https://doi.org/10.1056/NEJM197810262991705

Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., Dietterich, T. G., & Müller, K.-R. (2021). A unifying review of deep and shallow anomaly detection. Proceedings of the IEEE, 109(5), 756–795. https://doi.org/10.1109/JPROC.2021.3052449

Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. Neural Computation, 13(7), 1443–1471. https://doi.org/10.1162/089976601750264965

Shafique, A., Mehmood, A., & Elhadef, M. (2021). Detecting signal spoofing attack in UAVs using machine learning models. IEEE Access, 9, 93803–93815. https://doi.org/10.1109/ACCESS.2021.3089847

Subramanian, N., Suresh, K., & Logesh, R. (2023). A comprehensive review of smart attendance monitoring systems. Sensors, 23(6), 3041. https://doi.org/10.3390/s23063041

Zeng, J., Cao, B., & Tian, R. (2020). Quality monitoring for micro resistance spot welding with class-imbalanced data based on anomaly detection. Applied Sciences, 10(12), 4204. https://doi.org/10.3390/app10124204

Zhu, Y., Xie, J., Liu, J., & Gao, X. (2022). Fraud detection in employee attendance records using machine learning. Applied Sciences, 12(7), 3517. https://doi.org/10.3390/app12073517