Scientific journal of Research studies in Future Electrical Engineering, 2025, Vol. 2, No. 1, pp. 48-55

Scientific journal of Research studies

in Future Electrical Engineering

www.journalhi.com/pow/

Predictive Risk Prioritization for Telecommunication Site Projects Using Network Health Indicators and Operational Data

Salar SeyedAhmadi [1,*]

1- Ph.D. in Telecommunication Engineering (Fields and Waves), Urmia University, Urmia, Iran. Email: salar.seyedahmadi@gmail.com.

* Corresponding Author

Abstract

Telecommunication site projects operate within complex technical environments in which network performance, equipment reliability, and operational conditions can influence the effectiveness of installation, integration, and commissioning activities. Conventional monitoring approaches often rely on predefined thresholds or isolated indicators, which may not adequately represent interactions among network measurements or distinguish abnormal conditions according to their potential operational significance. This study develops a methodological framework for predictive risk prioritization using network health indicators and operational data, with particular emphasis on the distinction between network anomaly prediction and project-level risk assessment. The proposed framework consists of five stages: dataset verification, temporal feature extraction, predictive model development, performance evaluation, and risk prioritization. Publicly available telecommunications observability data, particularly the TelecomTS dataset, are identified as potential sources for examining multivariate network behavior and anomaly patterns. Logistic regression, random forest, and gradient boosting are specified as candidate predictive algorithms. The evaluation protocol incorporates precision, recall, F1-score, receiver operating characteristic area under the curve, precision-recall analysis, and ranking-oriented measures. Temporal or grouped data partitioning is proposed to reduce information leakage and provide a more credible assessment of generalization to previously unseen observations. A central methodological contribution is the separation of predicted anomaly probability from operational severity. The proposed prioritization score combines these components to support the relative ranking of events requiring further investigation. However, the score is treated as a candidate decision-support measure rather than a validated estimate of project failure, financial loss, or schedule delay. The reviewed literature supports the feasibility of machine-learning-based anomaly detection in telecommunications networks, while also highlighting the importance of data quality, validation design, and operating conditions. The study establishes a reproducible analytical structure for connecting network observability with risk-oriented decision support. Nevertheless, the proposed models have not yet been independently executed on the selected dataset, and no new comparative performance results are claimed. Furthermore, the available public network data do not establish direct relationships with project-level outcomes. Future empirical validation using linked network telemetry, maintenance records, commissioning results, and project schedules is therefore necessary to determine whether the framework improves operational prioritization in actual telecommunication site projects.

Keywords: Telecommunication Site Projects; Predictive Risk Prioritization; Network Health Indicators; Anomaly Detection; Machine Learning

1- Introduction

Telecommunication networks increasingly depend on the coordinated operation of radio access equipment, transmission infrastructure, software-based network functions, and supporting power and environmental systems. As network architectures become more heterogeneous and operational measurements grow in volume, identifying emerging problems at individual sites becomes increasingly challenging. A site may maintain acceptable average performance while exhibiting localized signs of instability, whereas an isolated deviation may reflect a temporary fluctuation rather than a persistent technical problem. Distinguishing between these conditions is important for maintenance planning, network reliability, and the management of telecommunication infrastructure projects.

Network observability provides opportunities to examine operational conditions through multiple measurements rather than relying exclusively on individual performance indicators. Mobile network operators collect numerous parameters and key performance indicators to evaluate service quality and identify conditions that may affect network operation. However, the number of variables involved and the uncertain distributions of their measurements complicate the detection and interpretation of abnormal behavior. Ramírez et al. [1] addressed this challenge by proposing an explainable machine-learning methodology for detecting and classifying performance anomalies in mobile networks. Their research demonstrates the relevance of combining data preparation with interpretable analytical models when examining complex network behavior.

Temporal analysis provides another important perspective. Network measurements may evolve gradually before a disruption becomes apparent, while other anomalies may emerge abruptly. Methods capable of learning patterns across successive observations can therefore complement conventional threshold-based monitoring. Oğuz and Kalaycıoğlu [2] investigated long short-term memory networks and one-dimensional convolutional neural networks for detecting anomalous conditions in densely deployed cellular networks. Their

study considered signal-related indicators, including signal-to-interference-plus-noise ratio and channel quality information, to distinguish healthy cells from degraded or anomalous states. The findings support the relevance of temporal modeling for cellular network monitoring, although the reported performance is specific to the study's experimental setting.

The use of multiple indicators is particularly relevant when a network condition does not manifest uniformly across all protocol layers. Frey et al. [3] demonstrated that an experimentally introduced abnormal encryption state in an LTE radio access network could be detected through measurements collected at a different layer from the one where the change originated. Their study used measurements from 1,987 experimental runs and 285 variables, illustrating how information distributed across network layers can contribute to anomaly detection. The experiment also emphasized that performance and generalization depend on the control and consistency of the evaluation environment.

These studies provide a methodological foundation for examining network health through multivariate and temporal observations. Nevertheless, identifying an anomaly is not equivalent to determining its operational priority. Monitoring systems may generate multiple alerts without establishing which condition deserves immediate investigation, particularly when technical resources must be distributed across numerous sites and concurrent deployment activities. A risk-oriented approach must therefore consider not only whether an abnormal condition is likely to occur, but also how its significance should be evaluated in the operational context.

Predictive modeling offers a means of estimating the likelihood of abnormal network conditions from historical observations. However, the operational usefulness of such predictions depends on the way they are translated into decisions. An anomaly detector may identify several deviations during the same monitoring period, but these deviations may differ substantially in persistence, affected components, and potential consequences. A brief fluctuation in a noncritical indicator may require observation, whereas a persistent abnormal condition affecting a component necessary for commissioning may warrant immediate investigation.

Gorshenin et al. [4] proposed a probability-informed machine-learning approach for analyzing mobile network traffic and identifying network-related events. Their research illustrates how probabilistic modeling can complement conventional traffic analysis and support the examination of network behavior over time. Nevertheless, probabilistic anomaly assessment and project risk estimation remain distinct analytical tasks. The probability of an abnormal network condition cannot automatically be interpreted as the probability of installation delay, commissioning failure, or additional project expenditure.

This distinction is important for telecommunication site projects, where technical implementation is influenced by interdependent activities, including equipment installation, transmission integration, power readiness, testing, and commissioning. Network measurements may provide useful evidence about operational conditions, but project-level consequences must be represented through appropriate outcome variables. Without verified project records, a model can estimate the likelihood of a defined network anomaly but cannot establish its predictive accuracy for project delays or financial losses.

The availability of public telecommunications observability datasets creates an opportunity to investigate this problem using reproducible analytical procedures. Feng et al. [5] introduced TelecomTS, a multimodal dataset derived from a 5G telecommunications network, designed to support research on time-series analysis, anomaly detection, and related observability tasks. The dataset includes heterogeneous measurements and contextual information, providing a potential basis for evaluating methods that analyze multivariate network behavior. Its documented characteristics make it relevant to anomaly prediction, although the dataset does not independently establish a direct relationship between network anomalies and site-project outcomes.

Against this background, the present study develops a methodological framework for predictive risk prioritization using network health indicators and operational data. The framework separates the estimation of anomaly likelihood from the assessment of operational severity. It proposes a structured procedure for extracting temporal features, comparing predictive algorithms, evaluating model performance, and ranking events according to their estimated likelihood and defined severity. The empirical component is intended to assess network anomaly prediction using publicly accessible data, while the extension to project-level risk management is treated as a separate validation requirement.

The study has three objectives. First, it examines the suitability of multivariate network indicators for predicting abnormal operational conditions. Second, it establishes a comparative evaluation procedure for candidate machine-learning algorithms using appropriate classification and ranking metrics. Third, it formulates a transparent prioritization framework that distinguishes model-generated probabilities from independently defined severity criteria. These objectives provide the basis for the problem statement and methodology presented in the following sections

2- Problem Statement

Telecommunication site projects depend on the coordinated completion of technical and operational activities, including equipment installation, power integration, transmission connectivity, radio configuration, testing, and commissioning. Disruptions affecting these activities may arise from equipment faults, abnormal network behavior, configuration errors, or unexpected changes in operational conditions. Although network monitoring systems provide substantial volumes of performance and observability data, translating these measurements into a reliable and actionable risk-prioritization process remains challenging.

A central difficulty is that anomaly detection and project risk prediction are not equivalent. An anomaly detection model identifies observations that deviate from learned or predefined patterns. Project risk prediction, by contrast, requires evidence connecting observed conditions to defined project outcomes, such as schedule slippage,

unsuccessful commissioning, repeated site visits, or cost overruns. Without linked project records, a model trained on network anomalies cannot reliably quantify the probability of these project outcomes.

A second challenge concerns the prioritization of multiple detected anomalies. Network operators and project teams may encounter several abnormal conditions simultaneously, but these conditions do not necessarily have the same operational significance. A useful framework should therefore distinguish the probability of an abnormal condition from its severity and from the evidence supporting its potential consequences. Combining these dimensions without a documented rationale could create misleading rankings, particularly when severity assessments are subjective or inconsistent.

A third challenge is the need for rigorous evaluation. Temporal dependence, class imbalance, differences between network environments, and potential information leakage can make predictive models appear more effective than they are in practice. Consequently, model comparison should use an appropriate validation design and multiple performance measures rather than relying on accuracy alone. Evaluation should also distinguish classification performance from the practical usefulness of a prioritized list of alerts.

The present study addresses these challenges by developing a reproducible framework for predictive prioritization of abnormal network conditions using multivariate health indicators and operational observations. Publicly accessible telecommunications data can support evaluation of the anomaly-prediction component. However, a claim that the framework predicts actual site-project risks requires additional outcome data linking network conditions to project records.

The study is guided by the following research questions:

1. To what extent can multivariate network health indicators support the prediction of documented anomalous operational conditions?
2. How should candidate machine-learning models be compared under a validation design that accounts for temporal structure and prevents information leakage?
3. How can estimated anomaly probabilities be combined with documented severity criteria to support transparent operational prioritization?
4. What additional data and validation procedures are required before anomaly-based rankings can be interpreted as predictions of site-project outcomes?

The scope of the study is therefore deliberately divided into two levels: network anomaly prediction, which can be evaluated using a suitable labeled observability dataset, and project-level risk prioritization, which requires separate validation against actual project outcomes. This distinction establishes the analytical boundaries of the research and prevents network-level predictive performance from being misrepresented as evidence of project-level predictive accuracy

3- Methodology

3.1. Research Design

This study adopts a quantitative, data-driven methodological design to investigate predictive prioritization of abnormal conditions in telecommunication networks. The research is structured around two analytically distinct components: predicting anomalous network states from operational indicators and translating validated predictions into a transparent prioritization framework.

The first component concerns the identification of abnormal operational observations. The second concerns the interpretation of those predictions in relation to operational severity and, where suitable records are available, site-project outcomes. Keeping these components separate is necessary because a high probability of a network anomaly does not automatically imply a high probability of project delay or commissioning failure.

The proposed methodology is designed to support reproducibility through explicit definitions of the prediction target, feature construction, data partitioning, model evaluation, and prioritization criteria. Telecom TS is considered a candidate primary dataset because it provides multivariate telecommunications observability data and documented anomaly information [5]. Its suitability for each experiment must be confirmed by inspecting the available fields, label definitions, and data structure before model training.

At the current manuscript stage, the methodology is specified, but the candidate models have not been independently executed. Accordingly, the procedures described below constitute the planned empirical protocol rather than a report of completed model experiments.

3.2. Data Sources and Analytical Scope

The primary candidate source is the publicly available Telecom TS dataset described by Feng et al. [5]. The dataset is relevant to time-series anomaly analysis in telecommunications because it contains heterogeneous observations associated with network observability. Before analysis, the data must be inspected to establish which measurements are available, how observations are organized, and whether anomaly labels are attached to individual time steps, time windows, or complete sequences.

The data preparation stage should document:

- The number of records and sequences actually available after loading the dataset.
- The operational variables and their data types.
- The definition and distribution of anomaly labels.
- The presence of missing values, duplicate records, and inconsistent measurements.
- The temporal structure and any available identifiers for grouping observations.
- The distinction between experimentally generated anomalies and anomalies observed under real operational conditions.

Published dataset characteristics should not be treated as a substitute for verification of the data actually loaded for analysis. In particular, the availability of anomaly labels does not establish the availability of site identifiers, project schedules, commissioning outcomes, or cost records.

The analytical scope is consequently defined in two stages. Stage one evaluates the prediction of labeled network anomalies. Stage two considers project-level prioritization only if independently verifiable project records can be linked to network observations using appropriate identifiers and time periods. If such records are unavailable, the study will report network anomaly prioritization and identify project-level validation as a limitation rather than claiming that actual project risks have been predicted.

3.3. Target Variable and Feature Engineering

The target variable will be derived from the dataset's documented anomaly labels. For a binary classification experiment, the target can be represented as:

$Y_i = 1$ if observation *i* is labeled anomalous; $Y_i = 0$ if observation *i* is labeled normal.

This formulation is appropriate only if the source labels support a defensible normal-versus-anomalous distinction. If the dataset distinguishes multiple anomaly categories, a multiclass analysis may be considered separately, provided that each category has sufficient observations and a consistent definition.

Feature engineering will focus on extracting meaningful characteristics from available operational measurements. Candidate features include the mean, standard deviation, range, temporal trend, and mean absolute change within a defined observation window. These features may help characterize level shifts, variability, and abrupt changes in network behavior.

For a measurement series $x_1, x_2, ..., x_T$, the mean is defined as:

$$\bar{x} = (1/T) \Sigma_{t=1}^{T} x_t$$

The standard deviation characterizes variation within the observation window:

$$s_x = \sqrt{[(1/(T-1)) \Sigma_{t=1}^{T} (x_t - \bar{x})^2]}$$

The mean absolute change is defined as:

$$D_x = (1/(T-1)) \Sigma_{t=2}^{T} |x_t - x_{t-1}|$$

These features will be calculated only for variables for which the relevant measurements are available and meaningful. Categorical variables will be encoded using a procedure appropriate to their documented semantics. Missing values will be handled through a strategy fitted exclusively on training data.

A critical methodological requirement is the prevention of target leakage. Features must not incorporate information that would only become available after the prediction time. Similarly, overlapping windows from the same underlying sequence should not be distributed across training and test sets in a manner that allows the model to exploit nearly duplicated information.

3.4. Data Partitioning and Validation

The dataset will be partitioned into training, validation, and test subsets. Where timestamps and sequence identifiers are available, the preferred approach is a chronological or group-aware split that reflects the intended prediction setting. The training subset will be used to estimate model parameters, the validation subset to select configurations, and the test subset to assess final performance.

The partitioning procedure must be selected after examining the dataset's structure. A simple random split may be inappropriate if neighboring time windows are highly correlated or if multiple records originate from the same underlying experimental sequence. In such cases, temporal or sequence-level grouping is necessary to reduce information leakage.

All preprocessing steps, including imputation, scaling, feature selection, and any resampling of imbalanced classes, must be fitted or performed using training data only. The validation and test sets must remain independent of these fitting procedures.

The final evaluation should report the partitioning strategy, the number of observations in each subset, class distributions, and any exclusions made during preprocessing. These details are necessary for reproducibility and for interpreting whether performance estimates are likely to generalize beyond the observed sequences.

3.5. Candidate Predictive Models

Three candidate models are proposed for comparison:

Logistic regression. This model provides a comparatively interpretable baseline and estimates the probability of an observation belonging to the anomalous class. Its coefficients can help characterize the direction of association between selected features and the predicted outcome, subject to the assumptions and limitations of the fitted model.

Random forest. This ensemble method combines predictions from multiple decision trees. It can represent nonlinear relationships and interactions among network indicators without requiring all feature effects to be linear.

Gradient boosting. This method builds an ensemble sequentially, with each stage intended to improve predictive performance relative to the current ensemble. It can capture complex relationships in structured feature data, although its performance depends on appropriate model configuration and validation.

Each model will be trained on the same eligible training data and evaluated against the same test observations. Hyperparameter selection will use the validation procedure rather than the test set. Model selection will consider predictive discrimination, the ability to identify anomalous cases, and the suitability of predicted probabilities for risk ranking.

The candidate list is not a claim that these algorithms have already been trained on Telecom TS. Their inclusion defines the planned comparison. Any unavailable fields, incompatible labels, computational limitations, or dataset-specific constraints discovered during implementation must be documented.

3.6. Predictive Risk Prioritization Framework

The proposed prioritization framework separates the estimated probability of an anomalous condition from its operational severity. Let p_i denote the model-estimated probability that observation i is anomalous. Let S_i denote a documented severity score, where the scoring criteria and range must be defined before the ranking is calculated.

If a defensible severity score is available, a candidate composite index is:

R_i = 100 * p_i * (S_i / S_max)

where \(S_{\max}\) is the maximum permitted severity score. The resulting index is intended to support relative prioritization; it is not automatically a calibrated probability of project failure, financial loss, or schedule delay.

The severity score must be based on documented criteria, such as the affected component's operational role, the persistence of the condition, the number of affected services, or verified consequences. These criteria should not be inferred solely from the anomaly probability, because doing so would combine the same evidence twice and could exaggerate the resulting priority.

If the dataset does not contain reliable severity information, the composite index will not be calculated as though such information were available. In that case, observations will be ranked by predicted anomaly probability alone, and the lack of a validated severity component will be reported explicitly.

A project-level risk index would require an additional outcome model or validation procedure based on observed project consequences. The proposed anomaly score cannot substitute for that evidence.

3.7. Evaluation Metrics

The models will be evaluated using multiple metrics because no single measure adequately captures all aspects of anomaly prediction.

- Precision: the proportion of observations predicted as anomalous that are actually labeled anomalous.
- Recall: the proportion of labeled anomalies correctly identified by the model.
- F1-score: the harmonic mean of precision and recall.
- Accuracy: the proportion of all observations classified correctly.
- ROC-AUC: the ability to discriminate between normal and anomalous observations across classification thresholds.
- PR-AUC: the area under the precision–recall curve, particularly informative when anomalous observations are relatively uncommon.

For operational prioritization, ranking-oriented evaluation should also be considered. Precision at *k*, for example, measures the proportion of the top *k* ranked observations that are genuinely anomalous. The value of *k* should be selected according to a clearly stated operational use case rather than chosen after inspecting test results.

Probability calibration should be examined if model outputs are to be interpreted as probabilities or used in a composite risk index. A model that ranks anomalies effectively may still produce poorly calibrated probability estimates. Accordingly, ranking quality and probability calibration should be reported as separate properties.

No metric values will be inserted until the models have been executed on a verified dataset and the evaluation results have been generated. Cross-study results from published literature will not be presented as if they were results from the current experiment.

3.8. Reproducibility and Research Integrity

The implementation should record the dataset version, preprocessing decisions, feature definitions, model configurations, partitioning strategy, and evaluation procedures. Random seeds should be documented wherever stochastic procedures are used. Final test results should be produced only after model and feature choices have been finalized.

The analysis must also distinguish between published dataset facts, independently calculated experimental findings, and proposed methodological components. This distinction is essential because the present framework addresses a plausible route from network observability to risk prioritization, but the validity of that route must be established empirically.

4- Results

4.1. Scope of the Results

The results section distinguishes between documented characteristics of the public telecommunications dataset, findings reported in previously published studies, and experimental results that would require independent model execution. This distinction is necessary to ensure that the evidence presented in the paper accurately reflects what has been verified.

The current analysis has not yet produced independently executed comparisons of logistic regression, random forest, and gradient boosting on the selected dataset. Therefore, no new classification scores, model rankings, or claims of predictive superiority are reported here. The results presented in this section are limited to documented dataset characteristics and their implications for the proposed evaluation framework.

4.2. Characteristics of the Primary Dataset

The Telecom TS publication describes a large-scale observability dataset derived from a 5G telecommunications network. Its heterogeneous measurements preserve absolute scale information and support several analytical tasks, including anomaly detection and root-cause analysis [5]. These properties make the dataset relevant to research on multivariate network behavior.

The following characteristics are reported in the dataset description available for this study. They should be

confirmed against the exact dataset version used in any subsequent implementation.

Table 1. Reported characteristics of the TelecomTS dataset

Characteristic	Reported description
Application domain	5G telecommunications network observability
Data structure	Multivariate time-series observations
Number of samples	32,000
Sequence length	128 time steps
Number of channels	18
Floating-point channels	10
Integer channels	6
Categorical channels	2
Anomaly coverage	11 reported anomaly types, including synthetic scenarios and a real over-the-air jamming scenario

Source: Dataset description associated with Telecom TS [5]. The counts above are reported dataset characteristics, not results independently calculated by the present study.

The dataset's combination of numeric and categorical measurements creates opportunities to examine different representations of network behavior. Numeric indicators may characterize changes in measurement magnitude or variability, while categorical fields may provide contextual information about operating states. However, the usefulness of each field depends on its documented meaning, availability at prediction time, and relationship to the target label.

The presence of multiple anomaly types also has methodological implications. A binary classification task may provide an initial assessment of whether a model can distinguish normal from anomalous observations. Nevertheless, performance on an aggregate anomaly label may conceal differences between individual anomaly categories. If label counts and definitions permit, category-level evaluation should be conducted as a supplementary analysis.

4.3. Evidence from Previously Published Studies

Published studies provide evidence that machine-learning and statistical methods can identify certain abnormal conditions in telecommunications networks. Their results offer methodological context, but they cannot be treated as experimental results from the present study.

Table 2. Relevant published evidence

Study	Evidence reported	Relevance to the present research
Ramírez et al. [1]	An explainable methodology for detecting and classifying mobile-network performance anomalies, evaluated using operational network measurements and a synthetic testbed	Supports explainability and structured preprocessing
Frey et al. [3]	An experimental LTE study using 1,987 runs and 285 measured variables to detect an anomalous encryption state	Demonstrates the potential value of cross-layer measurements under controlled conditions
Feng et al. [5]	A public 5G observability dataset supporting anomaly detection and other analytical tasks	Provides a potential basis for reproducible model evaluation

Sources: The original publications by Ramírez et al., Frey et al., and Feng et al.

These studies differ in their datasets, experimental designs, anomaly definitions, and evaluation procedures. Consequently, their findings do not establish that one candidate algorithm in the present study will outperform another. They instead support the feasibility of studying anomaly detection using multivariate measurements and emphasize the importance of dataset-specific validation.

4.4. Implications for Predictive Risk Prioritization

The documented evidence supports three methodological observations.

First, multivariate observability data provide a plausible basis for identifying abnormal network conditions. Nevertheless, the usefulness of individual indicators depends on their stability, measurement quality, and relevance to the anomaly being predicted.

Second, the performance of an anomaly detector must be evaluated under a validation procedure that reflects the structure of the data. The existence of a public dataset does not, by itself, demonstrate that a model generalizes to new sites, different network configurations, or future operating periods.

Third, anomaly detection and project-risk prioritization require different forms of evidence. A model may identify abnormal network behavior, but project-level prioritization requires documented information about severity and actual consequences. In particular, the current dataset description does not establish verified links between each network observation and outcomes such as site commissioning delays, project cost overruns, or additional site visits.

For these reasons, the results presently support the feasibility of the proposed research design, rather than a claim that the complete project-risk prediction framework has already been empirically validated.

4.5. Summary of Findings and Evidence Boundaries

The available published evidence establishes that telecommunications observability data can be used to investigate abnormal network behavior and that public data are available for developing reproducible evaluation procedures. It also identifies an important limitation: the available evidence does not establish that a model trained on the selected anomaly labels can predict actual telecommunication site-project outcomes.

The present study therefore does not report independently measured precision, recall, F1-score, ROC-AUC, PR-AUC, or comparative model rankings. These values must be calculated from an executed and documented experiment before they can be presented as findings. The next section can interpret the methodological implications

of the documented evidence, but it must not describe proposed experiments as completed results.

5- Discussion

5.1. Interpretation of the Available Evidence

The evidence reviewed in this study supports the use of multivariate network observability data to investigate abnormal operating conditions in telecommunications infrastructure. However, the value of such data for project management depends on the distinction between detecting technical deviations and estimating their consequences for project delivery.

The Telecom TS dataset provides a potential foundation for evaluating anomaly-prediction methods using heterogeneous time-series observations [5]. The previously published studies also demonstrate the relevance of explainable machine learning and controlled experimentation for network anomaly analysis [1,3]. These findings support the methodological rationale for comparing predictive models, but they do not establish the performance of the proposed models on the selected dataset or validate predictions of project outcomes.

A further consideration is the operational meaning of an anomaly score. A model may assign a high score to an observation because its measurements differ substantially from learned patterns. That score indicates statistical or model-defined abnormality, not necessarily urgency, severity, or expected financial loss. Operational interpretation therefore requires contextual information about the affected component, the persistence of the abnormal condition, and its relationship to service continuity or project activities.

5.2. Methodological Implications

Three methodological implications follow from the available evidence.

First, model evaluation should reflect the temporal and structural properties of the data. When observations from the same time series are strongly correlated, random partitioning may allow related observations to appear in both training and test sets. This can overstate generalization performance. Chronological or sequence-level validation is therefore preferable when supported by the dataset structure.

Second, evaluation should extend beyond overall classification accuracy. In an imbalanced dataset, a model can achieve apparently strong accuracy while failing to identify a substantial proportion of anomalous observations. Precision, recall, F1-score, ROC-AUC, and PR-AUC provide complementary information. If predictions are intended to support operational prioritization, precision among the highest-ranked alerts and probability calibration are also relevant.

Third, explainability should be considered alongside predictive performance. Ramírez et al. [1] demonstrate the relevance of interpretable approaches for identifying network conditions associated with performance anomalies. In a practical deployment, explanations may help technical teams understand which indicators contributed to a prediction and whether the result warrants further investigation. However, feature importance alone does not establish a causal relationship between an indicator and a network fault.

Recent research also emphasizes the value of interpretable and reproducible anomaly scoring. Franco-Valiente et al. [6] describe a modular pipeline for mobile-network anomaly detection that uses a learned reference representation and normalized anomaly scores to support comparisons between observations. This provides a relevant methodological example of severity-oriented anomaly ranking, but its scoring procedure should not be treated as validation of the composite index proposed in this study.

5.3. Practical Implications for Telecommunication Site Projects

The proposed framework could support a staged workflow for network operations and site-project teams.

1. Monitor: collect relevant network health indicators and operational measurements.
2. Detect: estimate the likelihood that an observation represents an anomalous condition.
3. Prioritize: combine the prediction with independently documented severity criteria, where available.
4. Investigate: review high-priority observations with technical and operational personnel.
5. Validate: compare predictions against verified maintenance, commissioning, and project records.

This workflow could help structure the review of multiple alerts and provide a transparent basis for allocating diagnostic resources. Its practical value, however, depends on data quality, reliable labels, appropriate validation, and the integration of relevant operational context.

For project management applications, additional records would be needed to determine whether particular network conditions precede or coincide with actual project consequences. Useful fields could include site identifiers, equipment and configuration records, timestamps, work-order histories, planned and actual commissioning dates, documented failures, repeat visits, and verified cost or schedule deviations. Data linkage would require consistent identifiers and temporal alignment.

Once these records are available, a separate project-level model could be developed to estimate a defined outcome, such as commissioning failure or schedule delay. Its performance would need to be evaluated against actual project outcomes rather than against anomaly labels alone.

5.4. Limitations

The current study has several limitations.

First, the proposed candidate models have not yet been independently executed on the selected dataset. Consequently, no comparative performance metrics or empirical claims of model superiority can be reported.

Second, Telecom TS is designed for telecommunications observability research, not specifically for evaluating the delivery performance of telecommunication site projects.

Its anomaly labels do not, by themselves, establish whether an observation resulted in installation delays, commissioning failures, or additional costs.

Third, performance measured on one dataset may not generalize to different network vendors, operating environments, geographic settings, or site configurations. External validation would be required before applying a trained model to operational decision-making.

Fourth, the proposed composite prioritization index depends on the availability of a defensible severity measure. If severity is assigned through subjective judgments or inconsistent rules, the resulting rankings may not reliably represent operational urgency.

Finally, the relationship between network anomalies and project outcomes may be affected by external factors, including access restrictions, equipment delivery, power availability, permitting, and workforce scheduling. These factors are not adequately represented by network telemetry alone and would need to be included in a project-level analysis.

5.5. Future Research

Future research should proceed in three stages. The first stage is to execute and document the proposed anomaly-prediction experiments using a verified dataset and a leakage-resistant validation procedure. The second stage is to test whether the resulting anomaly scores provide useful operational rankings when evaluated against documented severity criteria. The third stage is to link network observations with actual site-project records and evaluate project-level outcomes independently.

Such a staged design would establish whether the framework progresses from anomaly detection to operational prioritization and, ultimately, to project-risk prediction without treating these tasks as interchangeable.

6- Conclusion

This study proposes a methodological framework for predictive risk prioritization in telecommunication site projects using network health indicators and operational data. The framework connects multivariate anomaly prediction with a structured prioritization procedure while distinguishing network-level abnormalities from project-level risks.

The proposed methodology includes temporal feature extraction, comparison of candidate machine-learning models, validation designed to reduce information leakage, and evaluation using classification, ranking, and probability-calibration metrics. It also introduces a conditional prioritization index that combines predicted anomaly probability with independently documented severity. The index is intended to support transparent operational decisions and should only be applied when the required severity information is available and defensible.

The study's principal contribution is the explicit separation of three analytical tasks: detecting abnormal network behavior, prioritizing operational alerts, and predicting project outcomes. These tasks require different target variables and validation evidence. A model that accurately identifies network anomalies cannot automatically be assumed to predict commissioning failures, schedule delays, or cost overruns.

The current evidence supports the methodological rationale for the proposed framework, but independent model execution and project-level validation remain outstanding. Consequently, no new comparative performance results or claims of predictive superiority are established in this manuscript version. Future empirical work should evaluate the candidate models using a verified dataset, document their performance under an appropriate validation design, and link network observations with actual maintenance, commissioning, and project-management records.

Such validation would determine whether network anomaly predictions can provide useful early-warning information for telecommunication site projects and whether integrating operational severity improves the practical quality of risk prioritization.

References

[1] Ramírez JM, Díez F, Rojo P, Mancuso V, Fernández-Anta A. Explainable machine learning for performance anomaly detection and classification in mobile networks. Computer Communications. 2023;200:113–131.

[2] Oğuz HT, Kalaycıoğlu A. Anomaly detection in multi-tiered cellular networks using LSTM and 1D CNN. EURASIP Journal on Wireless Communications and Networking. 2022;2022:101.

[3] Frey M, Evans T, Folz A, Gregg M, Quimby J, Rezac JD. Anomalous state detection in radio access networks: A proof-of-concept. Journal of Network and Computer Applications. 2024;231:103979.

[4] Gorshenin A, Kozlovskaya A, Gorbunov S, Kochetkova I. Mobile network traffic analysis based on probability-informed machine learning approach. Computer Networks. 2024;247:110433.

[5] Feng A, Varvarigos A, Panitsas I, Fernandez D, Wei J, Guo Y, Chen J, Maatouk A, Tassiulas L, Ying Z. TelecomTS: A multi-modal observability dataset for time series and language analysis. In: Proceedings of the 43rd International Conference on Machine Learning. Proceedings of Machine Learning Research. 2026;306:30439–30469.

[6] Franco-Valiente JM, Calle-Cancho J, Cortés-Polo D, et al. An efficiency-driven anomaly detection pipeline for mobile networks. The Journal of Supercomputing. 2026;82:361.