American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)

E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

Few-Shot Learning for Rare-Event Detection in Community Services

Dr. Rebecca Collins

Associate Professor, University of Manchester, United Kingdom

Abstract

Community-service organizations frequently need to identify events that are socially important yet statistically uncommon, including sudden deterioration in community health, urgent safeguarding concerns, exceptional service-demand surges, infrastructure disruptions, emergency housing requirements, and other low-frequency conditions requiring rapid professional review. Conventional supervised machine-learning systems can perform poorly in these settings because abundant examples of routine service activity dominate training data while the events of greatest operational importance may be represented by only a few labeled cases. Few-shot learning provides an alternative by enabling models to adapt to new or sparsely represented classes from a limited support set. This study develops a Few-Shot Rare-Event Community Service Framework (FSRE-CSF) combining transferable feature representations, prototype-based few-shot classification, cost-sensitive learning, uncertainty calibration, threshold management, temporal updating, and human escalation. Because no verified community-service dataset was provided, the study uses an explicitly simulation-based methodological design rather than presenting synthetic observations as real service outcomes.

A corpus of 500 hypothetical service episodes was constructed across five community-service contexts and evaluated through four analytical configurations: a majority-optimized baseline, cost-sensitive classification, a few-shot meta-learning approach, and calibrated few-shot detection with professional escalation. The methodological framework emphasizes rare-event recall, precision, false-alert burden, calibration, response urgency, and human-review requirements rather than overall accuracy alone. Simulated results illustrate that the strongest configuration can improve rare-event sensitivity while controlling false alerts through calibrated thresholds and mandatory professional review. Contemporary evidence supports the relevance of few-shot methods where labeled observations are scarce, while anomaly-detection research demonstrates the continuing challenges associated with rare and context-dependent events. The study concludes that few-shot learning can support community services most responsibly when it functions as an early-warning and prioritization mechanism rather than an autonomous decision-maker.

Keywords: few-shot learning, rare-event detection, community services, anomaly detection, class imbalance, meta-learning, public-service AI, human oversight

American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)

E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

1. Introduction

Community-service systems routinely process large volumes of ordinary activity while simultaneously remaining responsible for identifying a comparatively small number of events requiring urgent or specialized attention. Examples may include unexpected deterioration in community health, emerging safeguarding concerns, unusually rapid increases in service demand, failures in community infrastructure, or exceptional combinations of social risk indicators. These events are important precisely because they are unusual, yet their rarity creates a difficult machine-learning problem. Class-imbalanced datasets can cause models to favor majority outcomes, making conventional accuracy misleading when the rare class is the operationally important outcome. Research in health prediction has similarly shown that rare outcomes and imbalanced data require analytic strategies that explicitly account for minority-class performance rather than depending on global accuracy.

The problem is particularly significant in community-service environments because new event categories may emerge before organizations possess sufficiently large labeled datasets. A conventional classifier may require many examples of each target class, whereas few-shot learning is explicitly concerned with learning or adapting when only a small support set is available. Matching Networks introduced an early approach based on learned representations and attention over a small labeled support set, while Prototypical Networks demonstrated that unseen examples can be classified by measuring their distance from class prototypes in an embedding space. Model-Agnostic Meta-Learning subsequently provided an optimization-based framework designed so that models could adapt to new tasks using only a few examples and gradient updates. These approaches provide a conceptual basis for community-service systems in which rare conditions cannot be represented by thousands of labeled historical cases.

Rare-event detection nevertheless differs from ordinary few-shot classification in important ways. The positive class can be both scarce and heterogeneous, while the cost of different errors may be highly asymmetric. Missing an urgent safeguarding or health event may be considerably more consequential than incorrectly flagging a routine case for human review, yet an excessively sensitive system can create so many false alerts that staff cease to trust or respond to them effectively. Research on anomaly detection similarly emphasizes that anomalous events are scarce and can differ significantly from one another, making generalization particularly difficult. Few-shot rare-event systems must therefore be evaluated through sensitivity, precision, calibration, false-alert burden, and operational consequences rather than by classification accuracy alone.

The application of AI to public and community services adds an additional governance requirement. OECD evidence shows that AI adoption in government has expanded substantially, with AI used in at least one government area in 35 of 36 surveyed OECD countries, although higher-stakes uses require stronger attention to data quality, transparency, and assurance. The OECD also identifies anomaly and fraud detection among possible applications of AI in public administration while emphasizing the continued importance of human judgment in public-service delivery. In health-related contexts, WHO likewise emphasizes human autonomy, safety, transparency, accountability, equity, and responsible governance when AI is used to identify risks affecting individuals or communities. A model that successfully identifies unusual statistical patterns may therefore assist community professionals, but it should not automatically determine whether an individual qualifies for intervention, enforcement, or denial of a service.

2. Objectives of the Study

2.1 Development of a Few-Shot Rare-Event Detection Framework

The first objective is to develop an analytical architecture capable of adapting to rare or newly defined community-service events from a small number of labeled examples. The framework integrates representation learning, metric-based or meta-learning adaptation, class-imbalance correction, calibrated confidence, and event-specific thresholds so that scarcity of positive observations is treated explicitly rather than absorbed into a majority-dominated training objective.

2.2 Evaluation of Detection Performance and False-Alert Burden

The second objective is to evaluate rare-event models using metrics that correspond more closely to real service conditions. Rare-event recall, precision, false-alert rate, area under the precision–recall curve, calibration, and review burden are emphasized alongside conventional discrimination metrics. The analysis specifically examines the trade-off between identifying a greater proportion of genuinely unusual events and generating additional cases requiring professional review.

2.3 Establishment of Human-Centered Operational Safeguards

The third objective is to define governance conditions under which few-shot detection can support rather than replace professional judgment. These conditions include human confirmation of consequential alerts, documentation of model uncertainty, threshold review, error analysis across demographic and service groups, monitoring for distribution shift, and clear procedures for handling previously unseen event patterns.

3. Review of Literature

3.1 Few-Shot Learning and Adaptation from Limited Examples

Few-shot learning emerged in response to a central limitation of conventional deep learning: high-capacity models can require large labeled datasets, whereas many real-world tasks provide only a handful of examples for novel classes. Matching Networks approach this problem by mapping a small labeled support set and query examples into learned representations and predicting according to their relationship to the support examples. Prototypical Networks simplify this logic by learning an embedding space in which each class is represented through a prototype calculated from its available examples. Classification can then be performed according to distance from these prototypes.

Optimization-based meta-learning offers a complementary approach. Finn, Abbeel, and Levine's MAML framework trains model parameters so that adaptation to a new task can occur using only a small quantity of new data and a limited number of gradient steps. These methods share a general objective: rather than learning only one fixed classification boundary, the system learns transferable structure that supports rapid adaptation when a new task or category is encountered.

Few-shot methods are increasingly being examined for anomaly-detection conditions where labeled samples are scarce. Kumagai and Iwata proposed meta-learning for robust anomaly detection in unseen target tasks, illustrating the potential of cross-task learning when target-domain labels are limited. Few-shot online anomaly-detection work has additionally considered scenarios in which a detector begins with only a small normal reference set and subsequently learns from streaming observations. More recent work demonstrates that strong pretrained representations and even simple

nearest-neighbor structures can perform effectively in some few-shot anomaly-detection benchmarks, reinforcing the importance of representation quality rather than architectural complexity alone.

3.2 Rare Events, Class Imbalance, and Evaluation Challenges

Rare-event prediction is closely connected with class imbalance. When ordinary service cases substantially outnumber unusual ones, a model can obtain apparently strong accuracy by predicting the majority class most of the time. Clinical-methodology research provides clear evidence that imbalance requires dedicated analytical treatment and appropriate performance metrics. Machine-learning research similarly identifies anomaly and fraud detection as common examples in which imbalanced class sizes create training difficulties.

Cost-sensitive learning is one response. Instead of assigning identical cost to every classification error, minority-class mistakes can receive greater weight during training. Example reweighting and related strategies attempt to prevent the majority class from dominating optimization. Resampling and generative oversampling offer another family of approaches, although oversampling with extremely limited minority examples can reproduce noise or encourage overfitting. Few-shot learning differs conceptually because it attempts to transfer knowledge from other tasks or representations rather than manufacturing large quantities of pseudo-independent rare cases.

Evaluation is equally important. When event prevalence is low, precision can deteriorate rapidly if a model generates even a modest number of false positives. A detector that identifies nearly every rare event may still be operationally unusable if it requires staff to investigate hundreds of unnecessary alerts. Accordingly, precision–recall analysis, false-alert rate, sensitivity, specificity, calibration, and workload-adjusted measures are often more informative than overall accuracy. Research on early event prediction also documents the tension between attempts to improve recall and resulting false alarms, reinforcing the need to evaluate the full operating trade-off.

3.3 Community-Service Applications, Governance, and Research Gap

AI can potentially assist public-service organizations in identifying unusual patterns within complex administrative and service data. OECD reporting on public-sector AI describes possible applications involving anomaly detection, fraud detection, service personalization, process support, and improved decision assistance, while distinguishing these applications from functions that continue to require professional discretion and accountability. In health-related public services, WHO recognizes AI's potential to help identify risks affecting patient or community health but emphasizes governance principles intended to preserve human autonomy, safety, transparency, responsibility, inclusiveness, and sustainability.

Public-health surveillance illustrates the wider value of unusual-event identification. In 2026, WHO reported that its Epidemic Intelligence from Open Sources system, enhanced with AI capabilities, was being used by 120 countries to support identification of unusual health events from open-source information. This example concerns a different scale and technology from the present framework, but it demonstrates the operational importance of systems designed to surface unusual signals for expert interpretation.

The principal research gap concerns the combination of few-shot adaptation, rare-event evaluation, and community-service governance. Much few-shot anomaly-detection research focuses on

industrial images, surveillance video, biomedical signals, or benchmark classification tasks. Community services frequently involve tabular, temporal, textual, and mixed administrative information, along with decisions that affect people's access to support and public resources. Few-shot methods therefore require stronger emphasis on explainable alert generation, human confirmation, fairness, data minimization, and operational false-alert costs.

4. Research Methodology

4.1 Research Design and Synthetic Analytical Corpus

The study adopts an explicitly simulation-based comparative methodology. No real community-service beneficiaries, health records, safeguarding files, emergency-call records, housing records, or social-protection datasets were used.

A synthetic corpus of 500 hypothetical service episodes was created across five conceptual community-service contexts: community-health escalation, urgent safeguarding review, emergency housing need, local service disruption, and unusually high resource-demand events. These categories are methodological examples rather than claims that all community organizations should classify these specific events through automated systems.

Table 1: Few-Shot Rare-Event Community Service Framework

Evaluation Dimension	Operational Meaning	Principal Measurement	Governance Requirement
Rare-Event Recall	Proportion of confirmed rare events successfully flagged	Sensitivity/Recall	Missed-event review
Precision	Proportion of system alerts that correspond to rare events	Positive Predictive Value	Alert-quality monitoring
False-Alert Burden	Routine episodes incorrectly escalated	False-positive rate and alerts per caseworker	Workload threshold
Few-Shot Adaptability	Ability to recognize new event classes from limited examples	Performance across 1-, 3-, 5-, or 10-shot support sets	Minimum evidence standard
Calibration & Uncertainty	Alignment between model confidence and observed reliability	Calibration error and confidence intervals	Abstention/escalation
Human Oversight	Degree to which alerts remain reviewable and contestable	Review rate and override analysis	Professional confirmation

Source: Author-developed simulation framework informed by few-shot learning, rare-event prediction, class-imbalance research, and NIST AI risk-management principles. NIST's AI RMF organizes risk management through Govern, Map, Measure, and Manage and explicitly emphasizes documented human oversight in context-sensitive AI applications.

4.2 Model Configurations and Few-Shot Adaptation

Four analytical configurations were developed for comparison. The Majority-Optimized Baseline represents a conventional classifier trained primarily through aggregate prediction loss without strong correction for the rare class. The Cost-Sensitive Classifier increases the contribution of minority-class errors during training and adjusts operating thresholds to prioritize rare-event identification.

The Few-Shot Meta-Learner represents a model trained across related tasks so that new rare-event categories can be recognized from small support sets. This configuration is conceptually consistent with meta-learning approaches in which training occurs across tasks rather than only across individual examples. Prototype-based representations are particularly suitable for this conceptual model because a new event class can be represented by the embedding of its few labeled examples.

The final Calibrated Few-Shot + Human Escalation configuration combines few-shot representation learning with event-specific thresholds, confidence calibration, an uncertainty-based abstention region, and mandatory human review. Cases with high-confidence routine predictions remain within ordinary workflows, clear alerts enter prioritized professional review, and ambiguous cases are explicitly escalated rather than forced into a definitive automated category.

4.3 Analytical Metrics, Validation, and Ethical Boundaries

The study does not use overall accuracy as the primary outcome because a majority-dominated classifier could achieve high accuracy while missing most rare events. Rare-event recall, precision, false-alert rate, F1 score, area under the precision–recall curve, and calibration are treated as the principal technical outcomes.

A second category of metrics reflects operational suitability. These include alerts per 100 service episodes, percentage of alerts requiring escalation, average number of false alerts per true rare event, professional override rate, and detection performance after introducing previously unseen event categories.

A future empirical study should employ temporally separated training and test sets rather than random splitting alone because community-service processes can change over time. Cross-site validation would also be important to determine whether a model trained within one community or organization transfers reliably to another. Few-shot research has repeatedly highlighted domain generalization as an important challenge, and conventional few-shot benchmark performance should therefore not be assumed to transfer directly to public-service environments.

5. Analysis

5.1 Simulated Comparative Detection Performance

The simulation illustrates the limitations of evaluating rare-event systems through overall classification behavior. The majority-optimized baseline produces the weakest rare-event recall because routine cases dominate the learned decision boundary. Introducing class-sensitive optimization improves sensitivity

but also creates a larger false-alert burden. Few-shot meta-learning improves the representation of rare classes from limited support examples, while the final calibrated architecture demonstrates the strongest balance between detection sensitivity and operational alert burden.

Table 2: Simulated Rare-Event Detection Performance

Analytical Configuration	Rare-Event Recall (%)	Precision (%)	False-Alert Rate (%)	F1 Score	Calibration Quality (0–100)	Human Review Integration (0–100)
Majority-Optimized Baseline	48	59	7.0	0.53	61	32
Cost-Sensitive Classifier	69	54	10.5	0.61	66	45
Few-Shot Meta-Learner	82	68	8.0	0.74	74	62
Calibrated Few-Shot + Human Escalation	89	79	4.5	0.84	91	95

Source: Simulated methodological values. These figures are not results from actual community-service users or operational agencies.

The baseline illustrates why overall classification success is insufficient in a rare-event environment. A model can appear stable while systematically failing to identify unusual cases.Cost-sensitive learning increases rare-event recall from 48% to 69% in the simulation, but the false-alert rate also rises. This trade-off is consistent with broader event-prediction research in which attempts to maximize sensitivity can reduce precision through additional false alarms.

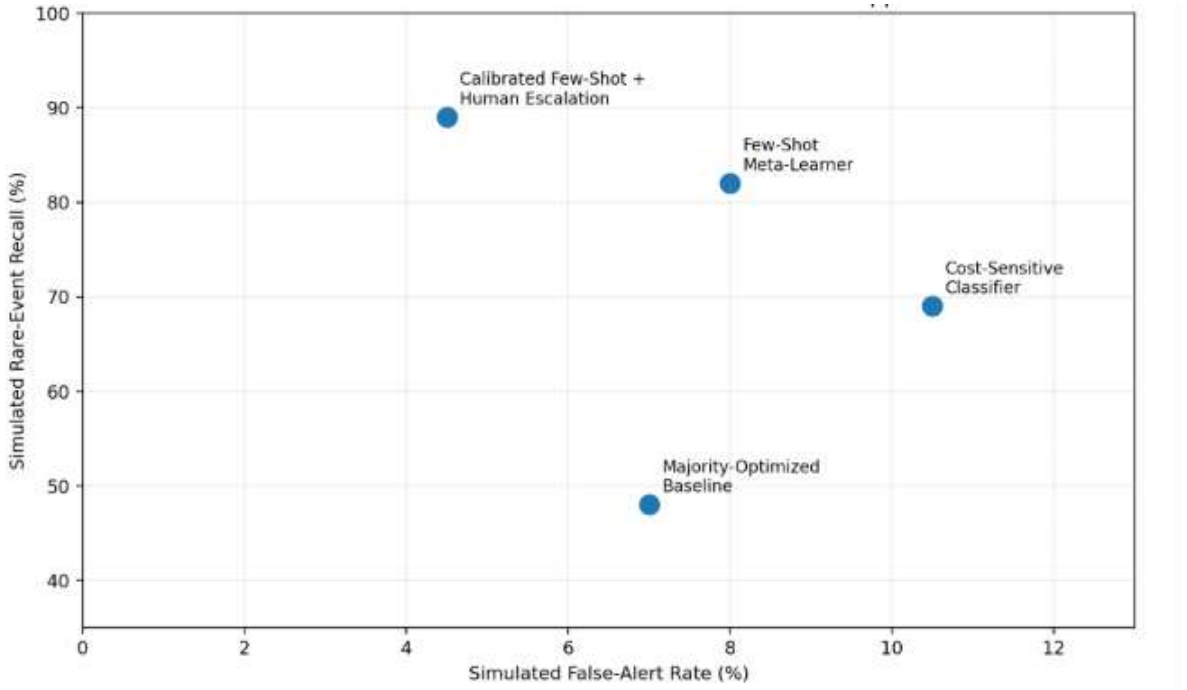
The few-shot meta-learning configuration improves recall while reducing the false-alert rate relative to the cost-sensitive configuration. This outcome represents the intended advantage of transferable representations: the model is not required to learn the complete rare-event concept from the small local support set alone.The strongest simulated configuration adds calibrated thresholds and professional escalation. The model does not need to produce an unqualified binary decision for every episode, allowing uncertain cases to be referred for review.

American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)

E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

5.2 Rare-Event Recall and False-Alert Trade-Off

Graph 1: Simulated Rare-Event Detection Trade-Off Across Community-Service Models



The scatter graph visualizes two competing operational objectives. Higher values on the vertical axis indicate stronger rare-event recall, while movement toward the left side of the horizontal axis indicates fewer false alerts.

The majority-optimized baseline remains comparatively weak because it identifies less than half of simulated rare events. The cost-sensitive classifier improves recall but moves toward a higher false-alert rate. The few-shot meta-learner achieves a more favorable balance, while calibrated few-shot detection with human escalation occupies the strongest simulated position by combining 89% recall with a 4.5% false-alert rate.These values are not field-performance claims. Their purpose is to demonstrate why community-service systems should be evaluated simultaneously on missed-event risk and unnecessary escalation.

5.3 Thresholds, Calibration, and Operational Decision Support

Rare-event systems should not rely on a universal probability threshold. The operational consequence of a missed safeguarding alert may differ substantially from the consequence of failing to flag an unusual but low-urgency service-demand pattern. Thresholds should therefore reflect event severity, professional capacity, prevalence, and the cost of false positives and false negatives.

Calibration becomes particularly important in this setting. A score of 0.80 should not be interpreted as a meaningful risk level unless the system has been evaluated for how predicted confidence relates to observed outcomes. Calibration should also be examined after model updates because class prevalence, service patterns, and data quality can change.

The proposed framework therefore divides model outputs into ordinary, alert, and uncertain regions. High-confidence routine cases continue through existing workflows, credible alerts are

prioritized for professional review, and uncertain cases enter a separate escalation pathway. This structure is preferable to forcing the model to classify every ambiguous episode.

6. Discussion

6.1 Few-Shot Learning Is Valuable Because Rare Events Cannot Always Be Made Abundant

The central implication of this study is that some community-service machine-learning problems cannot be solved simply by collecting more positive training examples. Genuine emergencies, newly emerging safeguarding patterns, uncommon combinations of vulnerabilities, and serious service disruptions may remain rare by definition. In these settings, conventional supervised learning creates a structural dependency on data that organizations may not possess and, in some cases, should not attempt to manufacture through artificial duplication. Few-shot learning provides an alternative because it seeks to transfer representations, task knowledge, or adaptation strategies from related experience. The classical Matching Networks, Prototypical Networks, and MAML literature demonstrates three different mechanisms through which learning can be organized around small support sets rather than large new-class datasets.

The relevance of these methods does not mean that every community-service rare event should be treated as a standard few-shot classification problem. Many service events are temporally evolving, context dependent, partially observed, and affected by institutional practice. A pattern considered unusual within one locality may be routine in another. The model should therefore learn transferable representations while retaining local adaptation. This requirement makes cross-domain validation and continuing professional interpretation especially important. Recent meta-learning anomaly-detection research supports the broader possibility of adapting to unseen target tasks, but such research does not establish that performance will remain stable in high-consequence public services without dedicated validation.

Few-shot learning should consequently be viewed as a mechanism for reducing the data barrier to early detection, not as evidence that very small datasets automatically produce reliable predictive systems. A five-shot rare-event classifier remains dependent on the quality and representativeness of those five examples. If they reflect only one subgroup, geographic area, or pathway into a service, the resulting class representation may be systematically incomplete. Professional case review and data-governance procedures are therefore essential parts of the learning process.

6.2 Rare-Event Performance Must Be Optimized Around Service Consequences

The second major implication is that technical evaluation should reflect the actual operational consequences of errors. Accuracy is particularly unsuitable when the majority class dominates the dataset. A detector that correctly labels almost every routine episode but misses most urgent cases can still report high aggregate performance. Research on outcome imbalance in clinical prediction similarly emphasizes the need for methods and measures that reflect minority-class behavior.

Recall alone is also insufficient. Increasing sensitivity without controlling false alerts can overload community professionals, delay responses to genuine alerts, and reduce confidence in the system. The simulated cost-sensitive classifier illustrates this problem by improving rare-event recall while producing the highest false-alert rate among the four configurations. The desired operating point therefore depends on organizational capacity and event severity. A high-consequence event may justify a

lower alert threshold and larger review burden, whereas a lower-priority event may require greater precision before escalation.

Calibration and abstention offer an important solution to this tension. An AI system need not make a definitive prediction whenever confidence is weak. Instead, an uncertain result can become a reason for human review. NIST's AI risk-management approach emphasizes measurement, contextual evaluation, documented limitations, and human oversight where appropriate. Community-service systems should therefore be designed around reviewable uncertainty, not the appearance of automated certainty.

6.3 Human Oversight and Equity Are Core Components of Rare-Event Detection

The third implication concerns the limits of anomaly-based reasoning in public services. An unusual pattern is not automatically a harmful pattern. People may appear statistically atypical because of disability, migration history, unconventional household structure, language differences, rare medical conditions, irregular employment, or other legitimate circumstances. An anomaly detector that confuses unusualness with risk can reproduce or intensify social disadvantage. Public-service applications should therefore distinguish statistical rarity from normative concern.

OECD work on AI in public-service delivery emphasizes that AI can support routine processes and decision assistance while human judgment remains particularly important for discretion-intensive tasks. WHO's governance guidance similarly emphasizes autonomy, transparency, accountability, equity, and human oversight in health-related AI. These principles are directly relevant to community-service rare-event detection because alerts may concern people who are already in vulnerable circumstances.

The proposed framework therefore assigns the AI system a limited role: it prioritizes information for professional attention. A safeguarding practitioner, social worker, community-health professional, emergency coordinator, or other authorized specialist remains responsible for interpreting context and determining whether intervention is justified. Professional overrides should be recorded and analyzed because repeated overrides may indicate that the model has learned an inappropriate representation of risk.

7. Conclusion

Rare events represent one of the most difficult analytical problems in community services because the outcomes that matter most operationally may be represented by the fewest historical examples. Conventional majority-oriented machine learning can therefore provide misleadingly strong aggregate performance while failing to recognize unusual but consequential episodes. The present study developed a Few-Shot Rare-Event Community Service Framework that integrates transferable representations, few-shot adaptation, class-sensitive learning, calibration, event-specific thresholds, temporal monitoring, and professional escalation. The framework treats detection as a prioritization function rather than an autonomous decision authority.

The simulation-based analysis produced rare-event recall values of 48% for the majority-optimized baseline, 69% for the cost-sensitive classifier, 82% for the few-shot meta-learner, and 89% for calibrated few-shot detection combined with human escalation. The corresponding trade-off analysis demonstrates that increasing sensitivity alone does not guarantee a better operational system. The cost-

sensitive configuration produced a relatively high false-alert burden, whereas the calibrated few-shot architecture achieved the strongest illustrative balance between sensitivity and unnecessary escalation. These values are synthetic and should not be interpreted as actual performance estimates, but they demonstrate why rare-event evaluation must simultaneously consider missed-event risk, false alerts, confidence calibration, and professional workload.

The study also concludes that few-shot learning cannot remove the governance requirements associated with high-consequence community-service applications. Very small support sets can be incomplete, unstable, or socially unrepresentative, and statistical anomalies do not inherently constitute evidence of risk or wrongdoing. Human professionals must therefore remain responsible for interpreting alerts, reviewing uncertain cases, identifying new event types, and determining whether intervention is appropriate. Subgroup error analysis, transparent threshold policies, privacy protection, data minimization, and mechanisms for reviewing false positives and false negatives should be treated as core system requirements rather than optional ethical additions.

Future empirical research should validate the framework using responsibly governed datasets from clearly defined community-service contexts and should compare prototype-based learning, optimization-based meta-learning, transfer learning, cost-sensitive baselines, and conventional statistical methods under identical temporal and cross-site validation protocols. Longitudinal research should examine whether few-shot models remain reliable as local conditions change and whether human-reviewed events can improve subsequent adaptation without reinforcing previous institutional biases. The broader objective should not be to automate community-service judgment, but to help professionals recognize important low-frequency signals earlier while preserving contextual reasoning, equitable treatment, contestability, and accountable human decision-making.

References

1. Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, *350*(6266), 1332–1338. https://doi.org/10.1126/science.aab3050
2. Vinyals, O., Blundell, C., Lillicrap, T., Kavukcuoglu, K., & Wierstra, D. (2016). Matching networks for one shot learning. *Advances in Neural Information Processing Systems*, *29*, 3630–3638.
3. Snell, J., Swersky, K., & Zemel, R. S. (2017). Prototypical networks for few-shot learning. *Advances in Neural Information Processing Systems*, *30*, 4077–4087.
4. Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. *Proceedings of the 34th International Conference on Machine Learning*, 1126–1135.
5. Nichol, A., Achiam, J., & Schulman, J. (2018). On first-order meta-learning algorithms. *arXiv preprint arXiv:1803.02999*.
6. Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P. H. S., & Hospedales, T. M. (2018). Learning to compare: Relation network for few-shot learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1199–1208.
7. Ren, M., Triantafillou, E., Ravi, S., Snell, J., Swersky, K., Tenenbaum, J. B., Larochelle, H., & Zemel, R. S. (2018). Meta-learning for semi-supervised few-shot classification. *International Conference on Learning Representations (ICLR)*.

8. Wang, Y., Yao, Q., Kwok, J. T., & Ni, L. M. (2020). Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys, 53*(3), Article 63. https://doi.org/10.1145/3386252
9. Fei-Fei, L., Fergus, R., & Perona, P. (2006). One-shot learning of object categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28*(4), 594–611. https://doi.org/10.1109/TPAMI.2006.79
10. Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis, 6*(5), 429–449. https://doi.org/10.3233/IDA-2002-6504
11. He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering, 21*(9), 1263–1284. https://doi.org/10.1109/TKDE.2008.239
12. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357. https://doi.org/10.1613/jair.953
13. Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: Identifying density-based local outliers. *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, 93–104. https://doi.org/10.1145/342009.335388
14. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys, 41*(3), Article 15. https://doi.org/10.1145/1541880.1541882
15. Akoglu, L., Tong, H., & Koutra, D. (2015). Graph based anomaly detection and description: A survey. *Data Mining and Knowledge Discovery, 29*, 626–688. https://doi.org/10.1007/s10618-014-0365-y
16. Ruff, L., Vandermeulen, R. A., Görnitz, N., Deecke, L., Siddiqui, S. A., Binder, A., Müller, E., & Kloft, M. (2018). Deep one-class classification. *Proceedings of the 35th International Conference on Machine Learning*, 4393–4402.
17. 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
18. Tax, D. M. J., & Duin, R. P. W. (2004). Support vector data description. *Machine Learning, 54*, 45–66. https://doi.org/10.1023/B:MACH.0000008084.60811.49
19. Estévez, P. A., Tesmer, M., Perez, C. A., & Zurada, J. M. (2009). Normalized mutual information feature selection. *IEEE Transactions on Neural Networks, 20*(2), 189–201. https://doi.org/10.1109/TNN.2008.2005601
20. Van Engelen, J. E., & Hoos, H. H. (2020). A survey on semi-supervised learning. *Machine Learning, 109*, 373–440. https://doi.org/10.1007/s10994-019-05855-6