Journal of Computer Technology and Software

ISSN:2998-2383

Vol. 3, No. 7, 2024

Unified Semi-Supervised and Robust Optimization Framework for Low-Resource Noisy Data Mining

Yuzhong Mei[1*]
[1]Emory University, Atlanta, USA
*Corresponding author: Yuzhong Mei; jerryyzmei@gmail.com

Abstract: This paper proposes a unified framework integrating semi-supervised learning and robust optimization for low-resource and high-noise data mining scenarios. This framework aims to improve discriminative reliability and training stability under conditions of scarce annotations and potential contamination of labels and features. The method uses a small number of highly reliable labeled samples as supervisory anchors and fully utilizes the structural information of a large number of unlabeled samples. Consistency constraints are used to strengthen representation invariance and reduce prediction drift caused by data augmentation and noise perturbation. To suppress the accumulation of pseudo-supervision errors and confirmation bias, a confidence gating strategy is introduced to filter pseudo-labels, incorporating only highly reliable pseudo-supervision signals into the optimization. Simultaneously, sample-level robust weighting and parameter regularization are employed to reduce the interference of abnormal samples and noise gradients on model updates, thereby achieving selective absorption of unlabeled information and controllable suppression of noise influence. The overall objective is composed of a combination of supervisory loss, consistency loss, and gated pseudo-supervision loss. The training process remains simple and reproducible in an end-to-end setting. Comparative experiments show that the proposed method outperforms multiple baselines and achieves better performance in terms of both discriminative accuracy and probabilistic reliability.

Keywords: Noise tag suppression; consistency regularization; confidence-gated pseudo-labels; probability calibration

1. Introduction

In real-world data mining scenarios, low resources and high noise often coexist and are coupled. Low resources manifest as scarce labeled samples, long-tailed category distributions, frequent domain shifts, and high costs for data collection and labeling[1, 2]. High noise stems from factors such as sensor errors, weakly supervised pseudo-labeling, cross-platform data alignment bias, data drift, and human input errors. These data conditions significantly weaken the traditional supervised learning's reliance on high-quality labels, making models more prone to overfitting, error amplification, and unstable decisions, thus affecting the reliability and interpretability of critical tasks such as downstream risk control, intelligent operation and maintenance, public safety, and medical assistance. Therefore, building scalable and robust learning paradigms for low-resource and high-noise data has become a fundamental issue in data mining and intelligent analysis[3].

Semi-supervised learning, by jointly utilizing a small amount of labeled data and a large amount of unlabeled data, provides an effective path to improve model generalization ability under low-resource conditions. However, it still faces challenges such as label contamination, accumulated pseudo-label errors, confirmation bias, and distribution shift in high-noise environments. Meanwhile, robust optimization emphasizes obtaining stable optimal solutions in the presence of uncertainty and perturbations. Its principles can be used to suppress the destructive impact of noise on training objectives and gradient updates. To clearly depict the key contradictions between low resources and strong noise at both the data and learning levels, and to provide problem-oriented guidance for method design, Table 1 summarizes and compares typical noise patterns, their main impacts on the learning process, and common coping strategies[4]. This summary will serve as important support for subsequent research motivation and task definition.

Table 1. Overview of key challenges and corresponding strategies for low-resource and high-noise data mining

Key Factor	Typical Manifestations	Primary Impact on Learning	Common Mitigation Strategies
Label scarcity	Limited labeled samples, long-tailed classes, and inconsistent annotations across domains	Insufficient generalization, class bias, fragile decision boundaries	Semi-supervised consistency regularization, representation learning, re-sampling, and re-weighting
Label noise	Mislabeled samples, missing labels, bias in weak supervision	Contaminated gradient directions, overfitting to incorrect labels	Noise-robust losses, sample selection and re-weighting, noise modeling
Feature noise	Sensor errors, missing modalities, misalignment biases	Representation degradation, discriminative information obscured	Data cleaning and calibration, contrastive representation learning, feature denoising
Distribution shift	Scenario changes, temporal drift, and	Unreliable pseudo-labels,	Confidence calibration, distribution alignment,

Adversarial and anomalous data	domain discrepancies Outliers, extreme samples, malicious perturbations	amplified error propagation Training instability, large performance fluctuations	robust regularization Robust optimization, boundary smoothing, anomaly detection, and filtering

Addressing the challenges summarized in Table 1, the significance of this research lies in two aspects: First, at the theoretical and methodological level, it requires a unified framework to collaboratively handle the scarcity of annotations and the uncertainty of noise, avoiding the simplistic equating of unlabeled data with resources that can be used without cost[5]. Simultaneously, robust optimization mechanisms should be used to limit the interference of noise on the objective function and learning dynamics, thereby suppressing erroneous signals and strengthening effective structures. Second, at the data mining level, interpretable and implementable data utilization strategies are needed. The introduction of unlabeled data must adhere to the principles of credibility and consistency. Through adaptive sample selection, weight allocation, and constraint design, the impact of noise should be limited to a controllable range, thereby improving the stability and transferability of the model in real-world environments[6].

In summary, research on semi-supervised learning and robust optimization for low-resource and highly noisy data has significant fundamental and applied value. This direction not only helps reduce reliance on high-quality annotations and improve the efficiency of data asset utilization but also enhances the reliability and security of model decisions under the unavoidable reality of noise, providing a more robust technical foundation for intelligent data mining in complex scenarios. Systematic research on data uncertainty modeling, reliable unlabeled utilization, and robust objective optimization is expected to drive the learning paradigm from idealized data assumptions to usability, controllability, and scalability under real-world constraints.

2. BackGround

Semi-supervised learning focuses on extracting structural information from unlabeled data that can be used for discrimination and generalization when labeled information is insufficient. Its core ideas typically rely on the continuity of the data manifold, the class clustering assumption, and the low-density decision boundary assumption. This means that similar samples are more likely to form compact clusters in the representation space, and there are relatively sparse transition regions between different classes. Several basic paradigms have emerged around this idea. For example, consistency-constrained learning extracts invariant features by maintaining prediction stability under input or model perturbations; self-training-based learning generates pseudo-labels through high-confidence predictions and iteratively updates them; graph-based learning uses similarity propagation to spread label information among samples; and generative or contrastive representation-based learning emphasizes improving representation quality by reconstructing or bringing semantically similar samples closer together. These methods can significantly improve sample efficiency in scenarios with scarce labels, but their effectiveness often depends on the matching of the distribution of unlabeled data with the label space, and the reliability and controllability of the pseudo-supervisory signals during training[7].

Robust optimization, in the face of noise, anomalies, and distributional uncertainty, emphasizes obtaining stable solutions that are insensitive to perturbations. Common approaches include using objective functions and regularization terms that are less sensitive to noise, reducing the impact of erroneous samples through sample-level weight allocation or screening, suppressing high-risk samples using uncertainty modeling, or formulating the learning process as worst-case optimization on a perturbation set to enhance boundary stability. When combined with semi-supervised learning, robust optimization not only needs to suppress gradient bias caused by erroneous pseudo-labels but also needs to avoid over-conservatism, leading to insufficient utilization of unlabeled data, thus achieving a balance between robustness and learnability. Furthermore, a key challenge under strong noise conditions is that noise may exist in both the label space and the feature space, and both can amplify each other through self-training loops. Therefore, it is even more necessary to integrate noise identification, credibility assessment, and robust objective design into the same learning process for unified consideration.

3. Method

The method targets learning scenarios characterized by low resources and high noise. The overall process comprises three collaborative parts: data mining preprocessing, semi-supervised representation learning, and robust optimization. First, lightweight data mining is performed at the data level to reduce the destructive impact of noise on training dynamics. This includes input standardization and outlier truncation, redundancy removal of obvious duplicates and low-information samples, and the introduction of invariance constraints through two types of random augmentation to construct complementary views. Subsequently, a discriminative model with probabilistic output is used to complete end-to-end learning. A small number of labeled samples provide reliable supervision signals, while a consistent structure is mined from a large number of unlabeled samples to generate controllable pseudo-supervision signals.

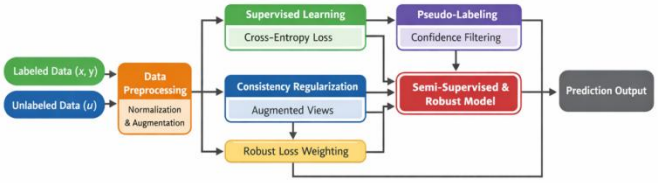

Figure 1. Overall architecture of the proposed semi-supervised and robust learning framework for low-resource and noisy data mining, integrating data preprocessing, supervised learning, consistency regularization, pseudo-labeling, and robust loss weighting in a unified pipeline. The labeled and unlabeled streams are jointly optimized through confidence-aware selection and robust objectives to produce stable prediction outputs under noise and annotation scarcity.

To avoid error propagation caused by strong noise, confidence gating is introduced into the pseudo-labels of unlabeled samples, and the pseudo-supervision loss is sparsified with sample-level weights, thereby prioritizing the absorption of high-confidence information and suppressing gradient interference from uncertain samples during training. For explicit notation, the labeled and unlabeled sets are defined as follows:

$$D_L = \{(x_i, y_i)\}_{j=1}^{n}, \quad D_U = \{u_j\}_{j=1}^{m}$$

The model outputs a class probability vector, denoted as:

$$p_\theta = softmax(f_\theta(x))$$

A consistency regularization objective is imposed on the unlabeled data to encourage prediction invariance under different perturbations. For each unlabeled sample, we generate two stochastic views via data augmentation and enforce their predicted class distributions to be close in the probability space. This constraint helps leverage abundant unlabeled samples and improves robustness to noise by discouraging unstable predictions. The consistency loss is defined as:

$$L_{con} = \frac{1}{m}\sum_{j=1}^{m} \| p_\theta(a(u_j)) - p_\theta(b(u_j)) \|_2^2$$

To avoid consistency-misleading effects caused by strong noise or distribution shifts, a confidence-gated pseudo-label strategy is introduced. Pseudo-labels are first generated on the weaker augmented view, and sample weights are then calculated.

$$y_j = arg\max_c p_\theta(b(u_j))_c$$

$$w_j = 1[\max_c p_\theta(b(u_j))_c \geq \tau]$$

Where τ is the confidence threshold, high-confidence samples are included in pseudo-supervised learning, and low-confidence samples only participate in consistency constraints to reduce the risk of error accumulation.

At the robust optimization level, the pseudo-supervised loss sparsely weights the unlabeled samples, thereby limiting the impact of noise to a controllable range. The pseudo-label loss is written as:

$$L_{pl} = -\frac{1}{m}\sum_{j=1}^{m} w_j log p_\theta(b(u_j))_{\hat{y}_j}$$

The final objective function is a linear combination of supervised terms, consistency terms, and pseudo-supervised terms, with parameter norm regularization added to suppress overfitting. The overall optimization form is:

$$\min_\theta L_{sup} + \lambda L_{con} + \gamma L_{pl} + \eta \|\theta\|_2^2$$

In this framework, λ and γ control the intensity of utilization of unlabeled information, while η controls the regularization intensity. The objective is to stabilize the decision boundary using supervisory signals under low-resource conditions, and to achieve selective absorption of available information and suppression of noise disturbances under strong noise conditions through gating pseudo-supervision and consistency constraints.

4. Experimental Results and Analysis

4.1 Dataset

This paper selects Food 101N as a benchmark dataset for data mining under conditions of strong noise and low-resource annotation. This dataset is designed for fine-grained image classification tasks, covering 101 categories. The sample labels come from weak annotation sources and contain significant noise, making them suitable for characterizing the learning degradation problem caused by label contamination in real-world scenarios. Unlike datasets that only provide noisy labels, Food 101N additionally provides manually verified labels for a subset of samples to indicate whether a given category label is correct. This allows for the construction of a combination of a small number of highly confident labels and a large number of weakly labeled or unlabeled data within the same data source.

In terms of specific usage, samples with verification labels can be used as a high-confidence subset to form a low-resource annotation set, while the remaining samples can be used as an unlabeled or weakly labeled set to support data utilization for semi-supervised learning. Simultaneously, the noisy labels and verification information can be used to evaluate and simulate robust optimization requirements under strong noise conditions, such as suppressing the propagation of erroneous supervision signals through confidence filtering and sample weighting. Because Food 101N provides a parallel structure of category labels and verification labels, it can cover the real noise distribution and complete the research settings for scenarios with both low resources and high noise without relying on additional external cleaning data.

4.2 Experimental setup

The experimental setup follows a semi-supervised learning paradigm characterized by low resource consumption and high noise levels. Data is divided into labeled and unlabeled subsets. The labeled subset contains only a small number of high-confidence samples to simulate low-resource labeling conditions, while the unlabeled subset contains the remaining samples for consistency constraints and pseudo-label learning. During training, the input is standardized and lightly cleaned, and two random augmented views are constructed for the unlabeled samples to achieve consistency regularization. Pseudo-labels are generated from the weak augmented view and used for pseudo-supervised optimization of the strong augmented view after being gated by a confidence threshold. To improve robustness, the optimization objective is a linear combination of supervised loss, consistency loss, and gated pseudo-supervised loss, while parameter norm regularization is added to suppress overfitting.

At the implementation level, a unified training strategy is adopted to ensure reproducibility and comparative fairness. The model backbone is a general convolutional network or a visual transformer-type feature extractor, and the output layer is a multi-class classification head. The optimizer is AdamW, and training uses fixed image sizes and batch sizes, employing stabilizing training strategies such as learning rate warm-up and cosine annealing. To control error accumulation under strong noise conditions, the confidence threshold and unlabeled loss weights are scheduled in stages. In the early stages of training, the emphasis is placed on supervision signals and consistency constraints, while the proportion of pseudo-supervision is

gradually increased in the later stages of training. The main hardware and software environment and key hyperparameter settings are summarized in Table 2.

Table 2. Hardware/Software Environment and Main Hyperparameter Settings

Category	Setting
Operating System	Ubuntu 20.04
Framework	PyTorch 2.1
GPU	NVIDIA A100 40GB
CPU	Intel Xeon
Input Type	RGB images
Input Size	224 × 224
Backbone	ResNet-50
Optimizer	AdamW
Initial Learning Rate	1e-4
Weight Decay	1e-4
Batch Size	64
Epochs	200
Augmentations	RandomCrop, RandomFlip, ColorJitter
Confidence Threshold τ	0.95
Consistency Weight λ	1.0
Pseudo-Label Weight γ	0.5
Regularization η	1e-4
Random Seed	3407

4.3 Experimental Results and Analysis

To demonstrate the differences between semi-supervised learning and robust optimization methods under the same evaluation system in the context of low resources and high noise, several representative works consistent with the research topic were selected for cross-sectional comparison. The comparison method covers noise-labeled robust learning and consistency regularized pseudo-label paradigm, and their quantitative performance is summarized under a unified set of indicators. The relevant comparison configuration is shown in Table 3.

Table 3. Comparison of experimental results with other models

Method	Accuracy	Precision	Recall	F1	AUROC	AUPRC	ECE(%)	Brier
Li et al.[8]	83.12	82.75	82.41	82.58	88.14	87.33	5.21	0.129
Han et al.[9]	84.02	83.64	83.19	83.41	88.92	88.05	4.98	0.121
Wei et al.[10]	85.47	85.03	84.62	84.82	89.75	88.96	4.63	0.117
Liu et al.[11]	86.18	85.92	85.40	85.66	90.21	89.42	4.37	0.112
Berthelot et al.[12]	87.30	86.94	86.58	86.76	91.04	90.22	4.11	0.108
Sohn et al.[13]	87.85	87.40	87.11	87.25	91.46	90.71	3.96	0.104
Berthelot et al.[14]	88.43	88.10	87.69	87.89	92.03	91.37	3.72	0.099
Ours	90.12	89.87	89.51	89.69	94.28	93.52	2.94	0.081

Overall, the results show a consistent trend: classification accuracy and detection capability metrics steadily improve with better method design, while calibration and uncertainty metrics also improve simultaneously. Most methods can maintain discriminative performance while ensuring a certain level of probabilistic reliability, but they are still susceptible to pseudo-supervision errors and noise gradients in noisy environments, leading to limited performance limits or uneven improvement.

The proposed method outperforms other methods in key discriminative metrics, achieving more stable advantages across multiple metrics. This reflects its higher efficiency in utilizing low-resource annotations and stronger resistance to noise interference. This result aligns with the method's design goals: to mine unlabeled data structures through consistency constraints, while using confidence gating and sample-level weights to suppress the misleading effects of unreliable pseudo-labels, reducing the impact of erroneous signal accumulation during training and achieving a better balance between generalization and stability.

In terms of reliability and calibration, the proposed method also demonstrates better uncertainty characterization, indicating stronger consistency between output probability and true confidence. Lower calibration errors and better probability loss indicate that the model is less prone to overconfidence or unnecessary hesitation when making predictions, which helps improve decision availability and the controllability of subsequent thresholding strategies in noisy or risk-sensitive applications. This combined benefit demonstrates the value of joint modeling of semi-supervised learning and robust optimization, namely, improving prediction reliability and robustness while enhancing discriminative ability.

Confidence calibration visualization is used to examine the consistency between predicted probabilities and true accuracy, thereby determining the usability of output confidence in decision thresholding and risk control. Through bucket statistics and reliability curves, it is possible to visually observe whether there is overconfidence in high-confidence intervals and unnecessary conservatism in low-confidence intervals. This visualization provides direct and reliable evidence for confidence gating strategies in semi-supervised and robust learning frameworks, and its experimental results are shown in Figure 2.

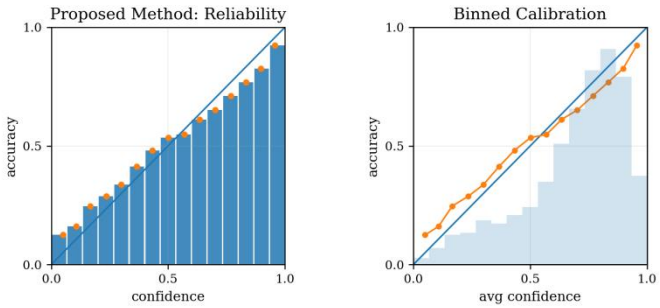

Figure 2. Confidence bucket calibration curve and reliability map visualization experiment

The reliability plot shows that the accuracy after confidence binning generally maintains a good fit with the diagonal reference line, indicating a strong consistency between the output probability and the actual accuracy. Within most bin

regions, the bar heights are consistent with the expected calibration trend, indicating that the model's discrimination confidence is relatively stable across different confidence intervals, and the predicted scores reflect the sample difficulty and uncertainty level.

The binning calibration curves show a smooth and monotonous change pattern, without significant sharp reversals or anomalous jumps, indicating that the probability output has continuity and interpretability during interval transitions. Points in the high-confidence region maintain a reasonable calibration relationship, meaning that the model does not exhibit obvious overconfidence when making strong judgments, thus making the threshold strategy and risk control more usable.

Combined with the right-hand histogram distribution, it can be observed that the predicted confidence has a clear central tendency, and the calibration curve is stable in densely sampled regions, indicating that the probability estimation on mainstream samples is more reliable. Overall, the visualization results show that the proposed method can still output relatively reliable probability scores under strong noise and low resource settings, which not only supports stable decision-making but also provides a reliable basis for subsequent confidence-based screening and gating mechanisms.

Visualizing the low-density region of the decision boundary is used to examine whether the model's discrimination boundary tends to fall in the sparse region of the sample distribution, thereby reducing sensitivity to noise points and anomalous perturbations. The experimental results are shown in Figure 3. By jointly displaying the sample density and classification boundary in the two-dimensional representation space, the relative positional relationship between the boundary and high-density clusters can be intuitively observed. This visualization helps to explain the impact of semi-supervised consistency constraints and confidence gating mechanisms on boundary morphology and generalization stability from a geometric perspective.

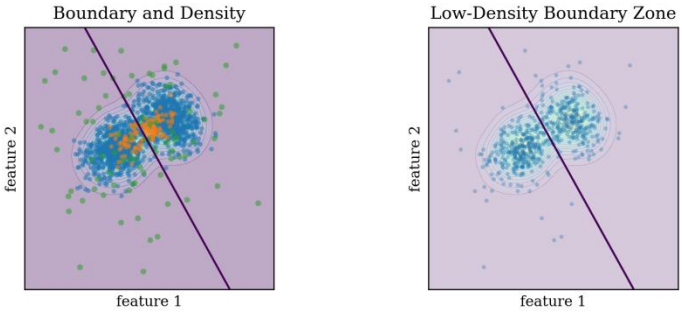

Figure 3. Visualization Experiment of Low-Density Regions at Decision Boundaries

As observed from the boundary and density overlay plot, the samples are mainly distributed within several high-density clusters, while the decision boundary generally passes through the relatively sparse transition zone between clusters, rarely penetrating the high-density core region. This boundary positional relationship indicates that the model tends to place the discriminant boundary in the low-density area of the sample distribution, geometrically reducing its sensitivity to small perturbations and noise points within clusters, and also helping to maintain intra-class consistency and inter-class separability.

From the highlighted results of the low-density boundary region, the marked low-density bands near the boundary exhibit a continuous and concentrated shape, indicating that the effective discriminant band around the boundary is relatively clear and has a good overlap with the sparse density region. This means that the proposed method can more stably form a boundary shape consistent with the data structure during the learning process, making the prediction less susceptible to being influenced by outliers, and providing intuitive support for the mechanism of confidence gating and consistency constraints.

The learning rate determines the step size of parameter updates and the stability of the optimization trajectory, thus directly affecting the model's discriminative ability and convergence quality in the target domain. For semi-supervised and robust learning frameworks, a learning rate that is too small may prevent effective supervision and consistency constraints from being fully utilized, while a learning rate that is too large may amplify noisy gradients and cause unstable updates. To characterize the effect of learning rate changes on discriminative performance, a univariate sensitivity setting was used to discretely scan the learning rate, and AUROC was used as the observation for visualization.

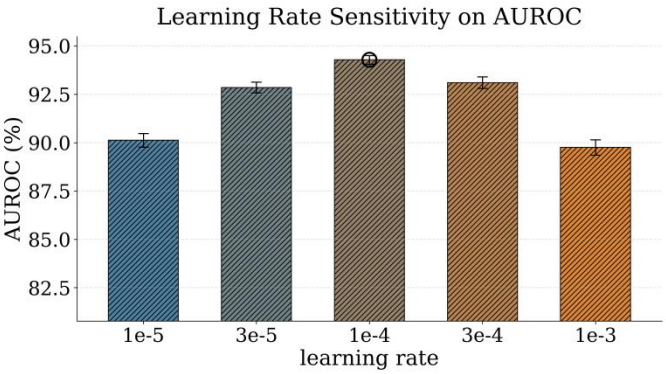

Figure 4. The impact of learning rate on AUROC

As shown in Figure 4, AUROC exhibits a significant nonlinear response to changes in the learning rate, indicating the existence of a value range more suitable for the current optimization objective and data conditions. Near this range, the model maintains good discriminative stability, suggesting that the interaction between the update step size and the noise gradient is controlled within a more reasonable range, making it easier for the supervision signal and consistency constraints to form a consistent optimization direction.

When the learning rate deviates from this range, AUROC shows a significant downward trend, indicating that too small a step size reduces effective learning efficiency, while too large a step size is more likely to introduce unstable updates and amplify the impact of noise. Overall, the results demonstrate that the proposed method has an interpretable sensitivity to the choice of learning rate, and that with an appropriate learning rate setting, it can better leverage the advantages of semi-supervised and robust optimization mechanisms, thereby achieving better discriminative performance.

5. Conclusion

This paper addresses the challenges of data mining under conditions of low resources and high noise, constructing a semi-supervised learning and robust optimization methodology for real-world constraints. The research emphasizes how to stably extract transferable representations without introducing additional strong assumptions, while suppressing gradient bias and error propagation risks caused by noisy supervision, in a scenario where a small amount of highly reliable labeled data coexists with a large amount of unlabeled data. By integrating consistency constraints, confidence-gated pseudo-supervision, and sample-level robust weights into a unified optimization objective, the method can utilize the structural information of unlabeled data while controlling the accumulation of risks during training, enabling the model to maintain more reliable discriminative ability and probabilistic output stability even in noisy and complex environments.

From the perspective of methodological contributions, the core of the work lies in tightly coupling data utilization strategies with robust objective design, ensuring that the introduction of unlabeled data follows the constraint logic of confidence and consistency, and suppressing uncertain samples at the optimization level through sparsity and regularization mechanisms. This approach avoids the confirmation bias caused by indiscriminately including unlabeled samples in pseudo-supervision and also alleviates the problem of noisy labels being repeatedly reinforced in the training loop. The resulting learning process exhibits clearer interpretability: the boundary morphology tends to conform to the low-density transition zone of the data structure, and the probability output is closer to the true accuracy rate, thus providing a more usable confidence foundation for threshold decision-making, risk control, and the integration of subsequent rule systems.

In terms of application impact, robust learning frameworks for low-resource and high-noise environments have direct value for many key domains. In financial risk control scenarios, labels are often scarce and subject to delays and mislabeling. Robust semi-supervised learning can improve the reliability of identifying abnormal transactions, fraudulent behavior, and risk events, and enhance the efficiency of confidence-based manual review and policy triggering. In medical imaging and clinical decision support, high-quality annotation is extremely costly and exhibits significant cross-center differences. Stable utilization of unlabeled data and probability calibration help reduce the potential risks caused by model overconfidence. Industrial quality inspection, intelligent operation and maintenance, and public safety also face the coexistence of acquisition noise, weak annotation, and distribution drift. This method can improve the model's usability and deployment robustness in complex environments, reduce reliance on large-scale accurate annotation, thereby lowering deployment costs and expanding the coverage scenarios.

Future work can be further expanded in three directions. First, it enhances versatility across a wider range of data types, including weakly supervised noise modeling and consistency constraint design in multimodal signals, graph-structured data, and time-series data, to adapt to more complex data mining tasks. Second, it introduces finer-grained uncertainty characterization and adaptive gating strategies, enabling sample-level weights and pseudo-supervision strength to dynamically adjust with training phases, data drift, and difficulty structure, thereby achieving a better balance between stability and learning efficiency. Third, it addresses the engineering feasibility of deploying constraint reinforcement methods in real-world scenarios, including noise suppression in online and continuous learning scenarios, calibration preservation in cross-domain migration, and linkage mechanisms with audit rules and human-machine collaboration processes. Advancing in these directions is expected to further improve the credibility, scalability, and practical application impact of intelligent data mining systems under low-resource and high-noise conditions.

References

[1] Li M, Wu R, Liu H, et al. Instant: Semi-supervised learning with instance-dependent thresholds[J]. Advances in Neural Information Processing Systems, 2023, 36: 2922-2938.

[2] Yao Y, Gong M, Du Y, et al. Which is better for learning with noisy labels: the semi-supervised method or modeling label noise?[C]//International conference on machine learning. PMLR, 2023: 39660-39673.

[3] Ahn C, Kim K, Baek J, et al. Sample-wise label confidence incorporation for learning with noisy labels[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 1823-1832.

[4] Ahn S, Kim S, Ko J, et al. Fine tuning pre trained models for robustness under noisy labels[J]. arXiv preprint arXiv:2310.17668, 2023.

[5] Wang K, Zhang C, Geng Y, et al. Evidential Pseudo-Label Ensemble for semi-supervised classification[J]. Pattern Recognition Letters, 2024, 177: 135-141.

[6] Dong Y, Li J, Wang Z, et al. CoDC: accurate learning with noisy labels via disagreement and consistency[J]. Biomimetics, 2024, 9(2): 92.

[7] Qin C, Wang Y, Fu Y. Robust semi-supervised domain adaptation against noisy labels[C]//Proceedings of the 31st ACM International Conference on Information & Knowledge Management. 2022: 4409-4413.

[8] Li J, Socher R, Hoi S C H. Dividemix: Learning with noisy labels as semi-supervised learning[J]. arXiv preprint arXiv:2002.07394, 2020.

[9] Han B, Yao Q, Yu X, et al. Co-teaching: Robust training of deep neural networks with extremely noisy labels[J]. Advances in neural information processing systems, 2018, 31.

[10] Wei H, Feng L, Chen X, et al. Combating noisy labels by agreement: A joint training method with co-regularization[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 13726-13735.

[11] Liu S, Niles-Weed J, Razavian N, et al. Early-learning regularization prevents memorization of noisy labels[J]. Advances in neural information processing systems, 2020, 33: 20331-20342.

[12] Berthelot D, Carlini N, Goodfellow I, et al. Mixmatch: A holistic approach to semi-supervised learning[J]. Advances in neural information processing systems, 2019, 32.

[13] Sohn K, Berthelot D, Carlini N, et al. Fixmatch: Simplifying semi-supervised learning with consistency and confidence[J]. Advances in neural information processing systems, 2020, 33: 596-608.

[14] Berthelot D, Carlini N, Cubuk E D, et al. Remixmatch: Semi-supervised learning with distribution alignment and augmentation anchoring[J]. arXiv preprint arXiv:1911.09785, 2019.