



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 15, Issue 2, February 2026**

# **A Review on Parameter-Efficient Continual Learning Frameworks for Scientific Text Classification**

**Kushal Sharma<sup>1</sup>, Akrshita Sharma<sup>1</sup>, Anamika Rangra<sup>2</sup>**

P.G. Student, Department of Computer Science and Engineering, Himachal Pradesh Technical University, Hamirpur,  
Himachal Pradesh, India <sup>1</sup>

P.G. Student, Department of Computer Science and Engineering, Himachal Pradesh Technical University, Hamirpur,  
Himachal Pradesh, India <sup>1</sup>

Guest Faculty, Department of Computer Science and Engineering, Himachal Pradesh Technical University, Hamirpur,  
Himachal Pradesh, India <sup>2</sup>

**ABSTRACT:** The proliferation of the scientific literature has generated the need to have automated systems that can efficiently process and categorize huge amounts of textual data. Knowledge discovery, information retrieval and document management in various areas of research is vital in the field of scientific text classification. Conventional deep learning models need vast retraining with new tasks or domains which makes them very computationally expensive and catastrophically forgetful. Other recent developments in transformer-based models and parameter-efficient learning methods like knowledge distillation and lightweight models have made them more efficient in computation and scalable to larger models [6]. Also, the data augmentation techniques were extensively studied in order to improve the model generalization and low-resource performance [1], [7]. Continual learning strategies and domain-specific language models have also proved to be effective in maintaining the already learned knowledge and adapting it to the new task [9].

In this paper, the parameter-efficient continual learning methods applied to scientific text classification are thoroughly reviewed. Transformer-based models, domain-specific language model, knowledge distillation models, data augmentation models, and continual learning strategies are studied. An analysis of the existing methods in terms of computational efficiency, scalability, and performance is also given. The limitations of the research and future trends of the development of efficient and scalable scientific text classification systems are also detected in the paper.

**KEYWORDS:** Scientific Text Classification, Continual Learning, Parameter Efficient learning, Transformer models, Domain Adaptation, Natural Language Processing.

## **I. INTRODUCTION**

The massive increase in scientific literature has resulted in a serious demand of automated ways to organize and categorize huge amounts of textual information. Scientific text classification can help in extracting knowledge easily, categorizing documents, and accessing information on even the most varied fields of research e.g. biochemistry, psychology, and computer science. A more recent research has investigated different methods to better scientific text classification such as fine-tuning large language models and dataset expansion methods [3], [4].

Transformer-based language models like BERT and its variants have been one of the most important driving factors in the recent advancement of natural language processing. These models are able to train the contextual representations based on large corpora and are also able to perform well in downstream tasks such as text classification. Transformer based methods and knowledge distillation have also been used to increase model efficiency and scalability [6]. Nonetheless, traditional methods of fine-tuning demand the models to be re-trained to support new datasets or domains. The result of this process is great computational cost, scaling problems, and catastrophic forgetting, in which already acquired knowledge is forgotten again.

To circumvent these drawbacks, repeated learning has been introduced to allow models to learn new tasks sequentially whilst being able to maintain previously gained knowledge. Moreover, parameter-efficient learning methods promote computation efficiency and performance in domain specific applications through updating fewer model parameters during adaptation and hence computational efficiency.

Also, pre-trained language models and data augmentation methods of a particular domain have been created to enhance special-purpose performance. Backtranslation and paraphrasing are examples of data augmentation techniques that can make models more robust and generalize better in low-resource settings [1], [7], whilst domain-specific models can help models perform better in biomedical and scientific applications [5], [9]. Regardless of these progress, issues with the heterogeneity of the domains, complexity of computations and effective transfer of knowledge are unresolved research questions [8].

The current paper gives an in-depth overview of the parameter-efficient continual learning methods in scientific text classification and discusses the effectiveness, limitations, and research difficulties.

The key contributions of the review are:

- Extensive literature review of transformer-based models and domain-specific language models.
- Learning and continual learning methods Analysis of parameter-efficient learning and continual learning methods.
- Comparative analysis of the available methods.
- Selection of future research directions.

## II. LITERATURE SURVEY

This section provides an in-depth overview of available literature regarding the field of scientific text classification, data augmentation, transform-based models, language models domain-specific, and parameter-efficient ways of learning. The literature is discussed in terms of a recent research that is concerned with enhancing the efficiency, performance, and scalability of models.

**Radlinski et al. 2025 [1]** compared the methods of data augmentation to backtranslation and paraphrasing as to enhance the performance of emotion classification in the large language model era. The research paper compared the conventional methods of augmentation with the ones based on LLM and proved that paraphrasing techniques produce more varied and meaningful training samples. These findings demonstrated better classification outcomes and generalization of models but issues of semantic consistency and computational cost were found.

**Sciuz-Bertrand et al. 2024 [2]** studied the approaches to zero-shot prompting and few-shot fine-tuning of large language models on document image classification. Its authors demonstrated that the concept of zero-shot prompting allows models to do classification without further training, whereas few-shot fine-tuning is more accurate based on limited labelled data. The research pointed out how effective prompt-based learning was but also pointed out the weaknesses that it was sensitive to prompt design and required computational power.

A comparative study of fine-tuning large language models on scientific text classification was introduced by **Rostam and Kertesz 2024 [3]**. The authors compared various pre-trained models and showed that transformer model fine-tuning models are substantially superior to the classical machine learning methods due to the ability to learn contextual scientific text representations. The research allowed highlighting the role of choosing suitable model structures and fine-tuning plans.

In their article, **Rostam and Kertesz 2025 [4]** suggested a better scientific text classification system on the basis of dataset expansion and hard-voting ensemble classification systems. The paper demonstrated that dataset expansion enhances the training diversity whereas hard-voting enhances the accuracy of classification through making predictions of multiple models. The experimental performance showed a better performance than the single models.

**Jahan et al. 2025 [5]** compared the efficiency of cost-effective large language models in benchmark biomedical tasks. The authors have considered the trade-off between model performance and the cost of computation, and they

demonstrated that the lightweight models could attain a competitive performance with lower resource consumption. The paper emphasizes the significance of the efficient model architectures of domain-specific applications.

The article by **Sanh et al. 2019 [6]** describes the creation of a smaller version of the BERT model, the DistilBERT, which was trained with knowledge distillation methods. The model has the benefit of having fewer parameters and it has good inference speed without significantly compromising the performance of the original model. This paper shows the usefulness of parameter-efficient learning methods in decreasing complexity.

**Chai et al. 2025 [7]** has thoroughly surveyed text data augmentation techniques based on large language models. The researchers identified rule-based, neural, and prompt-based methods of augmentation as well as addressed their strengths and limitations. The authors have emphasized the need of data augmentation to enhance model robustness, generalization and performance in low-resource settings.

**Rostam and Kertesz 2025 [8]** undertook the systematic overview of the development of pre-trained language model-based domain-specific text classification. The paper has examined different transformer architectures and highlighted how domain adaptation and transfer learning can be used to enhance classification. The authors also determined the issues of the heterogeneity of domains and computational complexity.

The language model MediBioDeBERTa proposed by **Kim et al. 2023 [9]** was based on continuous learning and intermediate fine-tuning methods and was trained on biomedical text. The model was trained and tested on the biomedical corpora and performed better than the current methods in various tasks. The paper emphasizes the role of the continuous learning process in maintaining knowledge and enhancing the performance in the domain.

Based on the analyzed research, it is seen that transformer-based models can offer powerful contextual representation of text categories, whereas the data augmentation method can enhance generalization and resilience. The parameter-efficient learning approaches lower the computational cost, and domain-specific models increase the specialization in special tasks. Nevertheless, the current methods are associated with the problem of computational complexity, domain adaptation, and knowledge retention. The constraints show that there is a necessity to develop efficient and scalable parameter-efficient continual learning models of scientific text classification.

### III. COMPARISION OF EXISTING METHODS

Table 1: Comparison of Existing Approaches for Scientific Text Classification

Ref. No.	Paper / Approach	Method Technique Used	Key Contribution	Advantages	Limitations	Application Domain
[1]	Backtranslation and Paraphrasing in the LLM Era	Data augmentation using backtranslation and paraphrasing	Compared traditional and LLM-based augmentation for classification	Improves data diversity and model generalization	Semantic inconsistency, high computation cost	Emotion classification
[2]	Zero-Shot Prompting and Few-Shot Fine-Tuning for Document Classification	Zero-shot and few-shot learning with LLMs	Evaluates classification without large labeled datasets	Reduces training data requirement, flexible	Performance depends on prompt design	Document image classification
[3]	Fine-Tuning Large Language Models for Scientific Text Classification	Fine-tuning transformer models	Comparative analysis of LLMs for scientific	High classification accuracy, strong	High computational cost, full retraining	Scientific text classification

			classification	contextual learning	needed	
[4]	Advancing Scientific Text Classification with Dataset Expansion and Hard Voting	Dataset expansion + ensemble hard voting	Improves classification using ensemble methods	Better performance and robustness	Increased training complexity	Scientific literature classification
[5]	Cost-Efficient Large Language Models in Biomedical Tasks	Lightweight and efficient LLM evaluation	Evaluates performance vs cost trade-off	Reduced computational cost, efficient deployment	Slight performance reduction vs large models	Biomedical NLP tasks
[6]	DistilBERT: Smaller, Faster Transformer Model	Knowledge distillation from large models	Reduces model size while maintaining performance	Faster training, lower memory, efficient inference	Some performance loss vs full models	General NLP tasks
[7]	Text Data Augmentation for Large Language Models (Survey)	Comprehensive survey of augmentation methods	Reviews augmentation techniques and challenges	Improves understanding of augmentation strategies	No experimental implementation	NLP and LLM training
[8]	Advances in Pre-trained Language Models for Domain-Specific Text Classification	Systematic review of domain-specific models	Reviews transformer models for specialized domains	Better domain knowledge representation	Domain adaptation challenges	Domain-specific NLP
[9]	MediBioDeBERTa: Biomedical Model with Continuous Learning	Domain-specific model with continual learning	Combines biomedical modeling and incremental learning	Prevents catastrophic forgetting, domain adaptation	Complex training process	Biomedical text classification

Based on Table 1, the main methods of existing approaches are oriented towards data augmentation, transformer-based fine-tuning and domain specific models. Nonetheless, issues like computational complexity, catastrophic forgetting, and scalability have not been solved yet. This points to the importance of parameter-efficient continual learning algorithms of scientific text classification.

#### IV. PROBLEM DOMAIN

##### 1. Rapid Growth of Scientific Data:

It is a problem of constant growth in the number of scientific works that complicates organization, control and search of appropriate information effectively [3], [4], [8].

##### 2. High Computational Cost of Traditional Models:

In conventional forms fine-tuning of large language models, retraining on new data results in the high computation complexity and resource use [3], [4].

##### 3. Catastrophic Forgetting in Sequential Learning:

Incremental learning Incremental learning models that are trained to perform new tasks tend to forget the prior knowledge, which lowers the performance [9].

4. **Domain Heterogeneity:**  
Science documents have both domain-based terminology and semantic structures, which general-purpose models are incapable of managing effectively [8], [9].
5. **Limited Labeled Data in Specialized Domains:**  
Numerous science disciplines do not include enough annotated data, which is why data augmentation and few-shot learning techniques are encouraged [1], [2], [7].
6. **Model Efficiency and Scalability Issues:**  
Transformer based models of large size are memory and computationally intensive and can only be practically deployed to a small scale [5], [6].
7. **Need for Parameter-Efficient Learning Approaches:**  
The adaptation process of all the model parameters adds cost to training, which generates pressure on effective learning strategies [6], [9].
8. **Challenges in Knowledge Transfer Across Domains:**  
Knowledge transfer across scientific fields is a research problem that has not been addressed effectively [3], [8].

## V. COMPARISON AND ANALYSIS

**Table 2:** Comparative Analysis of Existing Scientific Text Classification Approaches

Ref. No.	Learning Paradigm	Parameter Efficiency	Knowledge Retention Ability	Scalability	Data Requirement	Computational Overhead	Research Gap Identified
[1]	Data Augmentation-Based Learning	Low	Not addressed	Moderate	High labeled data	High	Lack of semantic consistency and efficient augmentation strategies
[2]	Zero-Shot / Few-Shot Learning	High	Limited	High	Very low labeled data	Low	Dependence on prompt engineering and model sensitivity
[3]	Full Model Fine-Tuning	Low	Poor (retraining required)	Limited	Large labeled datasets	Very High	Inefficient parameter usage and catastrophic forgetting
[4]	Ensemble Learning with Dataset Expansion	Low	Moderate	Limited	Large datasets	Very High	High complexity and resource requirements
[5]	Cost-Efficient Language Models	Moderate	Limited	High	Moderate data	Low	Trade-off between efficiency and performance
[6]	Knowledge Distillation	High	Limited	High	Large pretraining data	Very Low	Reduced representation capability

[7]	Data Augmentation Survey	Not Applicable	Not addressed	Depends on method	Varies	Varies	Lack of unified augmentation framework
[8]	Domain-Specific Pre-trained Models	Moderate	Moderate	Moderate	Domain-specific data	Moderate	Limited cross-domain generalization
[9]	Continual Learning with Domain Adaptation	High	Strong	High	Incremental data	Moderate	Complex training and stability issues

## VI. CONCLUSION

This paper provided a detailed overview of parameter-efficient continual learning methods of scientific text classification. The research has examined different models based on transformers, data augmentation, domain-specific language model, knowledge distillation models, and continual learning models. It has been shown in the comparative analysis that traditional fine-tuning methods achieve high classification performance with high computational cost and low scalability. The methods of data augmentation enhance the generalization of models but create the problem of semantic consistency and the computational cost. Knowledge-distillation and cost-effective models ease the model complexity and resource demands, whereas domain-focused and ongoing learning methods enhance the knowledge retention and domain adaptation.

The review points out that none of the methods at the outset should be effective in terms of high accuracy, computational efficiency and effective knowledge preservation. Thus the combination of parameter-efficient learning and continuous learning methods offers a new avenue to the creation of scalable and efficient scientific text classification systems.

## VII. FUTURE WORK

The following directions may be adopted in future research in scientific text classification:

- The creation of integrated mechanisms that will combine parameter-effective learning and constant learning to achieve enhanced effectiveness and memory.
- Planning of resilient strategies to avoid disastrous forgetting in sequential learning conditions.
- Discovery of sophisticated domain adaptation methods in the management of mixed scientific data.
- Enhancement of data augmentation procedures to achieve the semantic consistency and data quality.
- Design of low weight transformers with a lower computational power and memory.
- Exploration of methods of transferring knowledge across domains to enhance the generalizations in scientific domains.
- Testing of scalable and resource-efficient models to be deployed in large repositories in the real-world.

## REFERENCES

- [1] Ł. Radliński, M. Guściora and J. Kocoń, “Backtranslation and Paraphrasing in the LLM Era? Comparing Data Augmentation Methods for Emotion Classification,” in Proc. International Conference on Computational Science, Springer, Cham, pp. 3–17, 2025.
- [2] A. Scius-Bertrand, M. Jungo, L. Vögtlin, J. M. Spat and A. Fischer, “Zero-Shot Prompting and Few-Shot Fine-Tuning: Revisiting Document Image Classification Using Large Language Models,” in Proc. International Conference on Pattern Recognition, Springer Nature Switzerland, Cham, pp. 152–166, Dec. 2024.
- [3] Z. R. K. Rostam and G. Kertész, “Fine-Tuning Large Language Models for Scientific Text Classification: A Comparative Study,” in Proc. IEEE 6th International Symposium on Logistics and Industrial Informatics (LINDI), pp. 233–238, Oct. 2024.

- [4] Z. R. K. Rostam and G. Kertész, “Advancing Scientific Text Classification: Fine-Tuned Models with Dataset Expansion and Hard-Voting,” in Proc. IEEE 23rd World Symposium on Applied Machine Intelligence and Informatics (SAMI), pp. 527–532, Jan. 2025.
- [5] I. Jahan, M. T. R. Laskar, C. Peng and J. Huang, “Evaluating the Effectiveness of Cost-Efficient Large Language Models in Benchmark Biomedical Tasks,” arXiv preprint arXiv:2507.14045, 2025.
- [6] V. Sanh, L. Debut, J. Chaumond and T. Wolf, “DistilBERT: A Distilled Version of BERT: Smaller, Faster, Cheaper and Lighter,” arXiv preprint arXiv:1910.01108, 2019.
- [7] Y. Chai, H. Xie and J. S. Qin, “Text Data Augmentation for Large Language Models: A Comprehensive Survey of Methods, Challenges, and Opportunities,” arXiv preprint arXiv:2501.18845, 2025.
- [8] Z. R. K. Rostam and G. Kertész, “Advances in Pre-trained Language Models for Domain-Specific Text Classification: A Systematic Review,” ACM Transactions on Intelligent Systems and Technology, vol. 16, no. 6, pp. 1–41, 2025.
- [9] E. Kim, Y. Jeong and M. S. Choi, “MediBioDeBERTa: Biomedical Language Model with Continuous Learning and Intermediate Fine-Tuning,” IEEE Access, vol. 11, pp. 141036–141044, 2023.



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details