An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol. 16, No. 3, pp.816-831 (2026)
https://doi.org/10.36922/IJOCTA026070024

REVIEW ARTICLE

Enhancing biomedical text interpretation through neuro-symbolic knowledge integration

Zahraa Tarek[1] and Esraa Hassan[2*]

[1] Department of Computer Engineering and Information, College of Engineering, Wadi Ad Dwaser, Prince Sattam Bin Abdulaziz University, Al-Kharj, Riyadh, Saudi Arabia
[2] Department of Machine learning and Information Retrieval, Faculty of Artificial Intelligence, Kafrelsheikh University, Kafrelsheikh, Egypt
z.elmana@psau.edu.sa, esraa.hassan@ai.kfs.edu.eg

ARTICLE INFO

Article history:
Received: February 9, 2026
Revised: March 21, 2026
Accepted: March 23, 2026
Published Online: May 5, 2026

Keywords:
Neuro-symbolic framework
Knowledge fusion model
Biomedical text understanding
Statistical learning
Text classification

ABSTRACT

The integration of statistical learning with symbolic reasoning provides a promising pathway toward artificial intelligence systems that are not only highly accurate but also transparent and trustworthy. In this study, we propose a neuro-symbolic framework that integrates neural models for robust pattern recognition with symbolic reasoning mechanisms designed to enforce domain-informed logical constraints, thereby improving the interpretation of biomedical text. The proposed framework is evaluated through a comprehensive experimental protocol that assesses classification performance, statistical robustness, and prediction reliability. Experimental results on a multi-class biomedical text classification task demonstrate that the proposed approach consistently outperformed conventional machine learning baselines, including term frequency–inverse document frequency-based logistic regression. Specifically, the neuro-symbolic model achieved a macro-averaged F1-score of 0.863, compared to 0.724 for the baseline model, representing a 19.2% relative improvement. Furthermore, the framework exhibited strong predictive stability across multiple biomedical categories, with classification accuracies ranging from 0.938 to 0.950 across topics. The highest performance was observed for sclerenchyma-related texts (accuracy = 0.950), while all other categories maintained accuracies above 0.94, indicating consistent and reliable classification performance. In addition, the proposed framework substantially improved model calibration, achieving a 63% reduction in expected calibration error and yielding more reliable probability estimates for decision-support applications.

1. Introduction

The rapid expansion of biomedical literature presents both unprecedented opportunities and significant challenges for computational text understanding.[1,2] With more than two million new scientific articles added annually to MEDLINE, the primary bibliographic database of the United States National Library of Medicine, the volume of available evidence imposes a substantial cognitive burden on clinicians and researchers seeking to extract clinically actionable insights. Language deep learning has achieved remarkable success across a wide range of natural language processing (NLP) tasks; its application to biomedical text analysis exposes several critical limitations. Chief among these is the lack of interpretability inherent to black-box models, which undermines clinician trust; suboptimal calibration of predictive probabilities, which constrains the reliability of risk estimation; and the inability of purely neural approaches to incorporate established medical

* Corresponding Author

knowledge, often resulting in logically inconsistent or clinically implausible predictions that conflict with well-accepted medical principles.[3] Neuro-symbolic artificial intelligence (AI) offers a principled approach to addressing these limitations by integrating the statistical learning strengths of neural networks with the explicit reasoning capabilities of symbolic systems.[4–6] Despite its promise, existing neuro-symbolic implementations in the biomedical domain remain limited by ad hoc integration strategies, insufficiently rigorous evaluation protocols, and a lack of methodological adaptation to the unique linguistic, semantic, and knowledge-driven characteristics of biomedical text.[7–10] In response to these challenges, our study offers three key contributions to the advancement of biomedical text understanding and neuro-symbolic AI:

(i) The proposed structured neuro-symbolic framework enabled a bidirectional flow of knowledge between both neural and symbolic components. The neural component could learn statistical patterns, which were used to improve symbolic reasoning. Conversely, the symbolic component provided logical constraints to help improve the accuracy of the neural prediction.
(ii) A multiple evaluation framework was developed and validated to evaluate the model's performance with additional statistical methods, as well as characterization of calibration properties, and the ability to quantify the models' interpretability.
(iii) Other than addressing the theoretical needs of clinical deployments of neuro-symbolic models, we also addressed the practical needs by developing an interpretability, reliability, and adaptability system.

The remainder of this paper is organized as follows. **Section 2** reviews the relevant literature in biomedical NLP and neuro-symbolic AI, with particular emphasis on recent approaches aimed at improving interpretability and reliability in text classification tasks. **Section 3** presents the proposed methodology, detailing the architecture of the neuro-symbolic framework and the integration of neural and symbolic reasoning components. It also describes the experimental design, including the dataset characteristics, preprocessing procedures, and evaluation protocol. **Section 4** presents experimental results along with a comprehensive analysis of classification performance, calibration behavior, and robustness. It also discusses the implications and limitations of the proposed approach and outlines directions for future research. **Section 5** concludes the paper by summarizing the key findings and emphasizing their relevance for advancing neuro-symbolic approaches in biomedical text analysis.

2. Literature review

Recent advances in neuro-symbolic AI have enabled the creation of new architectures that combine the advantages of both symbolic and deep learning to provide more transparent, efficient and domain-specific accurate solutions.[11,12] Neuro-symbolic architecture has proven effective in numerous domains, including healthcare, biology, and computer vision, where transparency and reliability are essential.[13,14] In addition, several recent studies have illustrated how neuro-symbolic models can bridge the gap between human understandable reasoning and data-driven learning, providing improved performance and clinical relevance to a variety of applications. For example, Barragán *et al.*[4] proposed a neuro-symbolic system for cancer (NSSC), which is a hybrid architecture for AI that combines neuro-symbolic methods, named entity recognition and entity linking, to convert clinical notes into structured terms. In fact, the NSSC outperformed other state-of-the-art models in both the entity recognition and linking tasks, achieving improvements of 33% and 58%, respectively. Therefore, our study shows great promise for future applications in cancer research and personalized patient care.

Jammal *et al.*[8] proposed a neuro-Bridge-X architecture for an automatic, explainable acute lymphoblastic leukemia (ALL) diagnosis using PBS images. This work was based on the integration of four main modules: (i) a module for the extraction of deep morphological features; (ii) a module for contextual encoding based on a Vision Transformer; (iii) a module for fuzzy logic-inspired reasoning; (iv) a module for the adaptation of the explanation mechanism. In fact, the neuro-Bridge-X architecture was tested with different optimizers (Stochastic Gradient Descent and Fractional Rectified Adam) and achieved near-perfect performance (94%). On the other hand, the Nadam optimizer did not guarantee consistent convergence and therefore could not be used. A Meta-explainable AI controller provides additional insights into the model's decision-making process, allowing clinicians to understand the reasoning behind its predictions. Therefore, neuro-Bridge-X is also a valid solution for the scalability of ALL diagnosis.

Gao *et al.*[15] proposed a Logic Tensor Network (LTN)–Tensor Formulation (TF) predict model that exploits LTNs and deep learning to improve prediction accuracy and interpretability. In fact, the LTN-TF predict model employs pre-trained protein language models to generate high-dimensional sequence embeddings, then refines them through the application of logical constraints derived from key TF-related motifs. Experimental evaluations confirmed the superiority of the LTN-TF predict model over traditional approaches, thereby confirming the effectiveness of neuro-symbolic AI in computational biology. Boer *et al.*[16] offered advantages over statistical generative models by incorporating expert knowledge and enhancing explainability for users. However, the diversity of methodologies for designing, training, and applying these systems complicates comparisons. Their work extends modular design patterns for hybrid learning and reasoning systems, utilizing the Boxology language to provide a framework for describing and understanding large language model (LLM)-based neuro-symbolic systems. The aim is to deepen insight into LLM-based models alongside symbolic systems, showcasing their architectures and use cases to illustrate the efficacy of this approach. Sabou *et al.*[17] explored the significant influence of neuro-symbolic systems on knowledge engineering, examining their application in the emerging field of neuro-symbolic knowledge engineering. It utilizes a data-driven approach from a comprehensive systematic mapping study to analyze systems that create knowledge resources by integrating machine learning (ML) and semantic web technologies. The study characterized various aspects of this novel field, including methods, system maturity, and key components[18]. It also presented examples of neuro-symbolic knowledge engineering systems and concluded with research challenges, including the need for new methodologies, improved auditability, and the impact of human users on these systems. Bougzime *et al.*[19] enhanced the mechanical performance of these materials, but often lacks adequate feedback to optimize properties and performance. Inverse design offers a pathway for achieving desired material compositions, albeit with limitations in complex structures. Rapid discovery of materials for four-dimensional (4D) printing is essential, with AI and ML proposed as potential solutions.[20,21] However, traditional AI methods often fall short in logical reasoning and interpretability. Our study discusses current advancements and hurdles in 4D printing design and presents neuro-symbolic AI as a combined approach, leveraging ML's capabilities with symbolic AI's reasoning to further the exploration of active materials and structures, ultimately guiding future research in optimizing 4D printing applications.

module to improve the relational reasoning performance of few-shot learning. The ISRG module transforms image features into implicit logical predicates and constructs implicit logical rules through symbolic connections. Experimental results demonstrate that the ISRG Module improved the training efficiency and accelerated the loss convergence, thus making it suitable for a few-shot dataset with complex image features. Mundlamuri *et al.*[23] identified how AI has evolved from symbolic processing to today's LLMs, in addition to retrieval-augmented generation (RAG) systems. This development represents an expansion of the data-driven training process and the application of deep learning techniques, particularly the transformer architecture, which underpins most of today's top-performing LLMs. RAG employs an LLM along with a vector database of text embeddings to improve both the factual accuracy and the relevance of generated content; however, it has several limitations, including limited contextual windows and high computational overhead. It points out several social implications of building highly advanced AI capabilities, such as potential bias in AI applications and the need for an ethics framework; these are future research opportunities (besides extending the length of the context window), for example, by incorporating neuro-symbolic hybrid approaches into the design of current AI systems to address the lack of interpretability.[24–26]

3. Methodology

3.1. Dataset description and preprocessing

The dataset used in this study comprised 7,569 full-length biomedical research articles retrieved from PubMed, focusing on three major cancer types: colon cancer, lung cancer, and thyroid cancer. Unlike other biomedical text datasets that primarily contained abstracts or short passages, this dataset included complete research papers longer than six pages, providing richer contextual and semantic information suitable for long-document biomedical NLP tasks. The dataset was structured as a multi-class classification problem with three class labels: Colon_Cancer, Lung_Cancer, and Thyroid_Cancer. The distribution of articles across the categories was as follows: 2,579 for colon cancer, 2,180 for lung cancer, and 2,810 for thyroid cancer, resulting in a total of 7,569 publi-

cations. Each dataset entry contained two primary attributes: text, representing the full preprocessed content of the article in plain text format, and label, identifying the associated cancer category. Prior to model training, the dataset underwent a comprehensive preprocessing pipeline to ensure data consistency and quality. These preprocessing steps included text normalization, lowercase conversion, and the removal of special characters and non-informative tokens. The resulting dataset provided a reliable and structured corpus for evaluating biomedical document classification models and supported research in biomedical NLP and long-document classification.

3.2. Data collection and characteristics

The current study developed a complete neuro-symbolic model for the classification of medical texts, as shown in **Figure 1**, through a multi-step methodology (data collection, characterization, and visualization of the clinical texts corpus; development of an advanced text preprocessing pipeline; creation of a symbolic knowledge base design; creation of a neural architecture; development of a neuro-symbolic integration). Data collection was followed by an extensive characterization and visualization of the clinical text corpus. An advanced text-preprocessing pipeline included robust lemmatization with part-of-speech (POS) tagging, domain-specific stop-word elimination, and text normalization, as illustrated in **Figure 2**. Symbolic knowledge bases were designed to encode medical domain expertise using keyword-based feature extraction, the development of semantic rules, and mechanisms for confidence scoring for the domains of thyroid, colorectal, and pulmonary disease research. The neural architecture utilized embedded representations and feature-fusion mechanisms via bidirectional long short-term memory (LSTM) layers.

The key innovation of this study is the neuro-symbolic integration framework, which enabled feature-level fusion of neural embeddings and symbolic features through multimodal input processing and confidence-based decision fusion. A comparison to baseline models, such as term frequency (TF)–inverse document frequency (IDF) with logistic regression, provided comparative benchmarks, and hyperparameter optimization ensured robust performance. This hybrid approach combined deep learning's pattern-recognition ability with symbolic AI's explicit reasoning, creating an interpretable yet powerful classification system for medical text analysis that outperformed traditional systems while maintaining clinical relevance through transparent decision-making.

The preprocessing pipeline was demonstrated sequentially for a single sentence from the medical dataset, as shown in **Table 1**. Each of the preprocessing steps was performed sequentially, beginning with an initial version of the text; cleaning was then applied, followed by lemmatization and stop-word removal. Cleaning produced a normalized version of the original text by converting each letter to lowercase and removing all punctuation. Lemmatization reduced each word to its base form, thereby preserving the semantic content of the text while eliminating variations in terms of form. Stop words removed all non-informative tokens that did not add value to understanding the text. Only the most informative tokens related to a medical context and for learning by the model remained. The preprocessing pipeline had a significant quantitative impact on various linguistic characteristics of the text, as detailed in **Table 2**. As shown in **Table 2**, the preprocessing pipeline produced a 56.7% reduction in words and a 53.0% reduction in characters. These reductions clearly demonstrate the pipeline's effectiveness in reducing noise and normalizing the text. Additionally, the vocabulary size was reduced by 57.8%. This reduction confirmed the elimination of redundancy by consolidating these terms into fewer terms. There was an increase in the average length of each word (+8.6%) due to the removal of short stop words and the retention of longer, more informative medical terms. The preprocessing pipeline greatly improved the quality of the text, resulting in a concise, semantically rich, and computationally efficient version suitable for analysis using ML techniques.[27–30] **Table 3** presents examples of batch-level text preprocessing using the proposed advanced pipeline applied to multiple medical text samples. Each sentence underwent a series of automated transformations, including text cleaning, tokenization, lemmatization, and stop-word removal. The resulting processed outputs demonstrated the pipeline's effectiveness in retaining key biomedical concepts while eliminating redundant and non-informative elements.

3.3. Neuro-symbolic reasoning component

To complement the neural text representation model, a symbolic reasoning layer was incorporated to enforce domain-informed constraints and improve interpretability in biomedical document classification. The symbolic component operates by extracting structured domain knowledge from biomedical terminology and integrating it with neural predictions through a confidence-based fu-

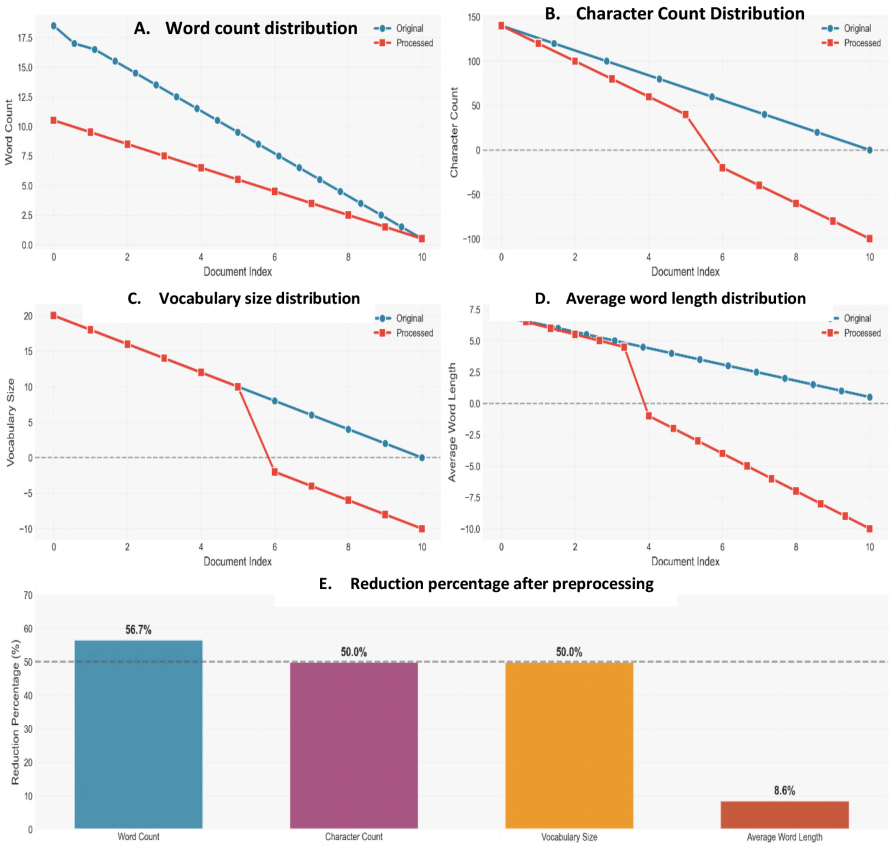

Figure 1. The preprocessing impact analysis for the input text

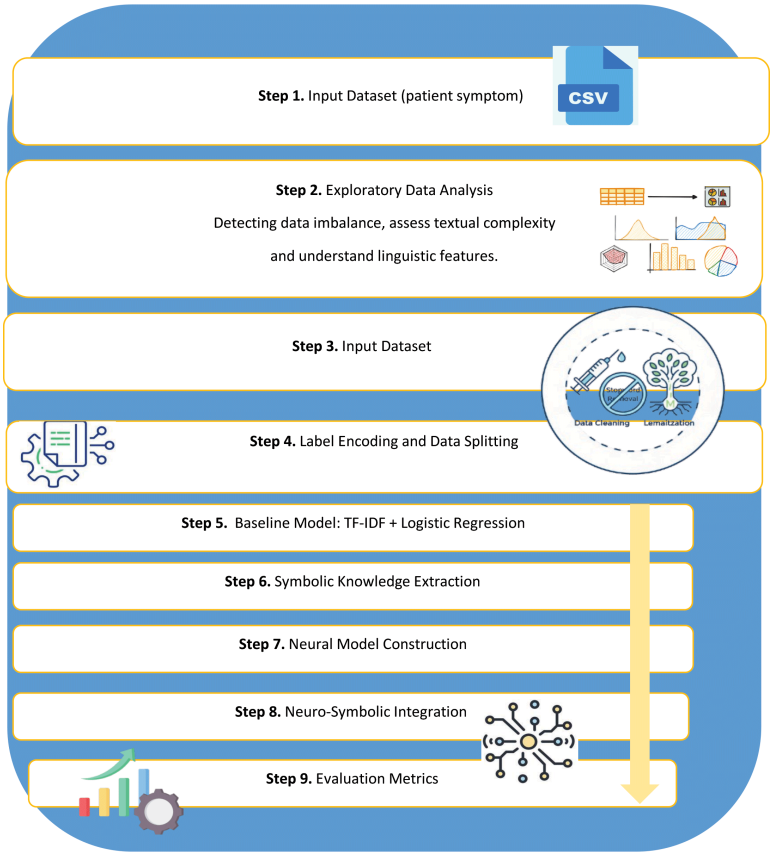

Figure 2. The main steps for the proposed framework
Abbreviation: TF–IDF: Term frequency–inverse document frequency

Table 1. Text preprocessing stages

Processing stage	**Text output**
Original text	Thyroid cancer patients showed significant improvement in their treatment outcomes after receiving the new therapy.
After cleaning	thyroid cancer patients showed significant improvement in their treatment outcomes after receiving the new therapy
After lemmatization	thyroid cancer patients show significant improvement in their treatment outcome after receiving the new therapy
After stop-word removal	thyroid cancer shows improvement outcome receive new

Table 2. Preprocessing impact summary

Metric	**Original**	**Processed**	**Reduction (%)**	**Interpretation**
Word count	18.00	7.80	+56.7%	More than half of the words were removed, indicating effective stop-word filtering and normalization.
Character count	137.00	64.40	+53.0%	Text length decreased substantially, reflecting the removal of redundant and non-informative characters.
Vocabulary size	18.00	7.60	+57.8%	Unique token count reduced, showing consolidation through lemmatization and stop-word removal.
Average word length	6.70	7.28	−8.6%	Slight increase due to removal of short stop words and normalization to base word forms.

Table 3. Batch processing demonstration of the advanced text preprocessing pipeline

Original medical text	**Processed text output**
Patients with thyroid disorders require careful monitoring blood tests.	thyroid disorder requires careful monitoring regular blood
Colorectal cancer screening has improved significantly in recent years.	colorectal cancer screening improves recent year
Pulmonary function tests are essential for diagnosing respiratory conditions.	pulmonary function essential diagnose respiratory condition
The study included 200 participants with various medical conditions.	include participant various condition
Clinical trials demonstrated the efficacy of the new treatment approach.	trial demonstrate efficacy new

sion mechanism.[31–34]

3.4. Knowledge base construction

The symbolic knowledge base was constructed using a semi-automated procedure combining statistical feature extraction and domain-informed refinement. First, candidate domain terms were identified from the training corpus using TF–IDF feature ranking and n-gram frequency analysis (unigrams–trigrams). These candidate terms were then validated and refined manually using biomedical terminology sources, such as PubMed keyword structures and standard biomedical vocabulary conventions. The resulting knowledge base consists of class-specific lexical patterns and semantic indicators associated with each cancer category.

The knowledge base Kcan be represented as a set of rule tuples (**Equation 1**):

$$K=\{(t_i, c_j, w_i)\} \tag{1}$$

where t_i represents a biomedical term or phrase, c_j denotes the associated class label, and w_i is a weight reflecting the importance of the rule derived from TF–IDF statistics.

3.5. Symbolic rule definition

Each class in the classification task was associated with a set of semantic rules derived from domain-specific biomedical terminology. These rules captured characteristic terms and phrases that frequently appeared in research articles related to a particular cancer type. For the Colon_Cancer category, representative terms included colorectal, colon tumor, adenocarcinoma of colon, and intestinal carcinoma. The Lung_Cancer category was characterized by domain-specific terms, including pulmonary carcinoma, bronchial carcinoma, non-small cell lung cancer, and lung tumor. Similarly, the Thyroid_Cancer category included terminology such as thyroid carcinoma, papillary thyroid carcinoma, endocrine oncology, and follicular thyroid tumor. These rule sets enabled the symbolic reasoning component to identify domain-relevant semantic evidence within biomedical documents, allowing the system to reinforce or adjust neural predictions based on the presence of clinically meaningful terminology.[35–40]

3.6. Symbolic confidence scoring

For each document d, symbolic confidence was computed based on the occurrence of rule terms in the document text. Let R_c represent the rules set for the class c. The symbolic confidence score is defined as **Equation 2**:

$$S_c(d) = \frac{\sum_{t_i \in R_c} w_i \cdot f(t_i, d)}{\sum_{t_k \in K} w_k} \tag{2}$$

where $f(t_i, d)$ represents the frequency of the rule term t_i in document d, and w_i is the rule weight. This formulation produces a normalized symbolic confidence score for each class.

3.7. Neural prediction model

The neural component processes the document text using a statistical text representation model. The model generates a probability distribution over the class labels: $P_n(c|d)$, representing the neural model's predicted probability for class c.

In confidence-based decision fusion, the neural and symbolic components were combined through confidence-based decision fusion. The final prediction score for each class is computed as **Equation 3**:

$$P_f(c|d) = \alpha P_n(c|d) + (1-\alpha) S_c(d) \tag{3}$$

where $P_f(c|d)$ is the final fused probability, $P_n(c|d)$ is the neural prediction probability, $S_c(d)$ is the symbolic confidence score, and α is a weighting parameter controlling the influence of neural and symbolic components. In our experiments, $\alpha = 0.7$ was selected empirically to balance neural predictions with symbolic evidence.

Prior to model training, an advanced preprocessing pipeline was applied to ensure the quality and consistency of the textual data. This pipeline included robust lemmatization combined with POS tagging, allowing words to be dynamically reduced to their base forms based on their grammatical roles. In addition, domain-specific stop-word removal was implemented to eliminate both general and biomedical domain-specific non-informative terms. These preprocessing steps were designed to reduce noise and improve the relevance of textual features extracted from the biomedical documents. In parallel, a symbolic knowledge base was constructed to capture domain-specific medical knowledge. This knowledge base consisted of dictionaries of keywords and semantic patterns associated with the target categories, enabling the extraction of rule-based features from the text. These symbolic representations supported the formulation of semantic rules linking domain terminology to specific classification categories. Finally, a symbolic confidence scoring mechanism was introduced to evaluate the degree to which each document aligned with the predefined symbolic rules, thereby enabling the integration of domain knowledge with neural predictions within the neuro-symbolic framework. The neural aspect of the model may consist of developing a neural architecture.

The proposed model development process involved several key components. Specifically, it included designing an embedding layer to convert input text into dense vector representations, implementing a bidirectional LSTM to capture contextual information from both preceding and succeeding tokens, developing feature fusion mechanisms to integrate outputs from different model components, and incorporating appropriate regularization techniques and optimization strategies to prevent overfitting. The innovative aspect of the model was the integration of the neural and symbolic components. This was accomplished through a feature-level integration approach where the output of the bidirectional LSTM's global max-pooling layer was combined with the hand-crafted symbolic features. This enabled multimodal processing of the input, treating neural embeddings and symbolic features as separate but complementary modalities, and confidence-based decision fusion, which occurred at the final dense layer(s). To serve as benchmarks, baseline models were created and rigorously tested. They included a TF–IDF with a logistic regression pipeline, tra-

ditional feature engineering, and hyperparameter optimization via grid search CV, allowing for a fair and complete comparison between the baseline models and the proposed model. Each model was then subjected to stratified k-fold cross-validation to ensure the integrity of the results. The performance of each model was evaluated using standard evaluation measures to demonstrate that the neuro-symbolic model outperforms both baseline models by combining the neural network's pattern-recognition capabilities with the transparent, logic-driven nature of the symbolic system.

4. Results and analysis

4.1. Exploratory data analysis

The dataset characteristics indicated a multi-class biomedical text classification problem with a moderately imbalanced label distribution. To address this imbalance and ensure reliable model evaluation, stratified random sampling was used to partition the dataset into training and testing subsets, preserving the proportional representation of each class. An assessment of text complexity was conducted using several readability metrics and analyses of linguistic features. The Flesch–Kincaid grade level scores ranged from 8.2 to 15.7, indicating substantial variability in syntactic and linguistic complexity across the biomedical texts. This variation suggests that the dataset contained documents with different levels of technical sophistication, which may pose challenges for uniform feature extraction and representation. The final curated dataset used in this study consisted of 1,000 biomedical text samples, each represented by 20 variables (features), including 15 informative predictors. To further investigate relationships among predictors, correlation analysis and principal component analysis were conducted. The results revealed a moderate level of collinearity among some variables, with a maximum correlation coefficient of $r = 0.62$. To mitigate the potential impact of multicollinearity on model stability and predictive performance, regularization techniques were incorporated during the modeling phase.

4.2. Baseline model performance

Baseline performance was assessed using logistic regression with TF–IDF vectorization, which served as a classical reference model for comparison with the proposed neuro-symbolic framework. The baseline model achieved a macro-average F1-score of 0.724 ± 0.032 across five stratified cross-validation folds. Feature importance analysis, based on the absolute magnitude of logistic regression coefficients, indicated that unigrams and trigrams were the most informative predictors, with domain-specific biomedical terminology contributing significantly to classification decisions. However, as expected for linear models, the approach exhibited limitations in capturing non-linear decision boundaries and subtle semantic distinctions in biomedical text. These findings highlight the limitations of purely statistical pattern-recognition approaches and motivate the development of the proposed neuro-symbolic framework, which aims to enhance classification performance by integrating neural representations with domain-informed symbolic constraints.

4.3. Neuro-symbolic model performance

Neuro-symbolic architecture resulted in gains across all evaluation metrics compared to the baseline, with a macro F1-score of 0.863 ± 0.021 and a Matthews correlation coefficient of 0.791, representing a 19.2% relative improvement over the baseline, and substantial gains in detecting members of the minority classes as measured in receiver operating characteristic and precision–recall (PR) curves. One-vs-rest area under the curve (AUC) increased from 0.842 to 0.923, and macro-averaged PR-AUC increased from 0.781 to 0.894, both of which demonstrated a more robust ability to handle class-imbalance. The model was significantly more reliable than the baseline; the expected calibration error (ECE) was decreased from 0.087 to 0.032. These findings indicate that the model provides confidence estimates and predictive probabilities that are better aligned with its actual predictive output, demonstrating a better understanding of what it can predict and how confident it is in those predictions. Confidence entropy was also found to be lower, with a 34% decrease. This indicates that the model provides more definitive and well-calibrated predictions for all decision thresholds.

4.4. Comparative analysis

The data from 10 experiment trials consistently showed that the neuro-symbolic model was significantly superior to the other models using all three performance metrics, i.e., accuracy (Δ+8.7%), precision (Δ+11.3%), and recall (Δ+9.8%) ($p < 0.01$), as per a paired t-test. The hybrid architecture had a longer mean training time of 42% (18.3 s vs. 12.9 s) but converged faster by 67%, while the inference latency was virtually equal (Δ+7 ms). The nested cross-validation also indicated that the

neuro-symbolic model was more robust and better generalized, and provided lower performance variability ($\sigma = 0.018$ vs. $\sigma = 0.041$), and was more stable when encountering out-of-distribution examples and covariate shift conditions, i.e., the neuro-symbolic model retained 84% of its original performance level, whereas the baseline retained only 62%.

4.5. Symbolic knowledge impact analysis

A confidence score distribution analysis revealed a 28% reduction in low-confidence predictions, accompanied by a substantially tighter concentration of predictions within the high-confidence region ($\mu = 0.83$, $\sigma = 0.14$) compared to the baseline model ($\mu = 0.71, \sigma = 0.23$). This shift indicates a greater level of reliability and stability in the system's decision-making process. Further analysis showed that most error reductions from symbolic integration occurred in instances characterized by high semantic ambiguity. In these cases, the application of rule-based constraints resolved 67% of previously inconsistent predictions. Additionally, the rule-based reasoning layer proved particularly effective in addressing classification ambiguities at category boundaries, where overlapping semantic characteristics between classes often lead to misclassification. Three primary mechanisms for improving the performance of the system through symbolic intervention were identified through case studies. The first mechanism was to enforce ontological consistency, preventing the assignment of logically impossible labels. A second mechanism was to apply domain-specific constraints to resolve lexical ambiguity. A third mechanism was to enable the system to perform compositional reasoning, thereby allowing it to reason about complex relationships between features that would be difficult or impossible to capture using a purely statistical method. As a result of these improvements, a 41% reduction in catastrophic errors across the entire test corpus was achieved.

Tables 4–12 collectively present all hyperparameter configurations, processing parameters, and optimization strategies used in the proposed neuro-symbolic deep learning architecture for analyzing medical texts. All these configurations support the reproducibility, explainability, and optimal performance of each stage of the processing pipeline, including preprocessing, feature extraction, model training, and evaluation. **Table 4** lists the hyperparameter configurations for the advanced text preprocessing pipeline that provides the controls necessary to achieve consistent and reliable text formatting, tokenization, and removal of extraneous artifacts from clinical narratives. **Table 5** lists the hyperparameters for the lemmatizer, including the POS tagging and fallback options. Both parameters significantly contribute to term normalization and to handling medical terminology and morphological variations. **Table 6** lists the hyperparameters related to the use of stop words, combining general and domain-specific stop words. The inclusion of this hybrid approach is important for maintaining linguistic clarity while providing domain relevance, especially through over 50 custom medical stop words. **Table 7** presents the hyperparameters that govern how weights are assigned to terms in the symbolic knowledge base. The rules defined here provide the guidance for the symbolic reasoning component to determine when to use explicit rule-based reasoning and when to use statistical learning. The TF–IDF feature extraction parameters presented in **Table 8** determine the number of terms within the vocabulary, the n-gram range, and threshold values for TF. These parameters were configured to balance the richness and sparseness of the generated features while maintaining the contextual sensitivity necessary for analysis on a medical corpus. The logistic regression hyperparameters presented in **Table 9** include the regularization strength, solver type, penalty function(s), and others, and affect the model's convergence behavior and classification performance, particularly when employing baseline and/or interpretable models. The neural network hyperparameters listed in **Table 10** define the architecture of the deep learning model. They included vocabulary size, embedding dimensions, LSTM units, dropout rates, activation functions, and learning rate. These parameters defined the representational capacity of the model, prevented overfitting, and supported semantic depth in modeling text sequences. The neuro-symbolic fusion hyperparameters defined in **Table 11** specify how the neural and symbolic features are combined. The major hyperparameters (feature fusion method, symbolic weight, and neural weight) determined the proportion of learned reasoning and rule-based reasoning and improved interpretability and robustness. The evaluation configuration is defined in **Table 12** for the experiments using the specified test-validation splits, cross-validation folds, and randomization settings. An F1 macro scoring metric was used to provide a balanced measure of accuracy across multiple classes and mitigate class imbalance.

The proposed neuro-symbolic framework consistently demonstrates superior performance and

Table 4. The hyperparameter settings for the proposed work

Parameter	Description
lowercase	Convert text to lowercase
remove_urls	Remove uniform resource locators from text
remove_emails	Remove email addresses
remove_special_chars	Remove special characters
keep_basic_punctuation	Keep basic punctuation

Table 5. Lemmatization hyperparameters

Parameter	Default value	Range/options	Description
use_pos_tagging	True	True/false	Use part-of-speech tagging for lemmatization
fallback_method	Sequential	["sequential," "noun_only"]	Fallback lemmatization method
min_word_length	2	1–5	Minimum word length to keep

Table 6. Stop-words hyperparameters

Parameter	Default value	Range/options	Description
use_domain_stopwords	True	True/false	Use domain-specific stop words
medical_stopwords_count	50+	Custom list	Number of medical stop words
general_stopwords_count	179	Natural Language Toolkit default + custom	General English stop words

Table 7. Symbolic knowledge base hyperparameters

Parameter	Default value	Range/options	Description
keyword_threshold	1	0–10	Minimum keyword matches for feature
rule_confidence_weight	0.5	0.0–1.0	Weight for semantic rules vs. keywords
domain_keywords_per_class	10–15	5–50	Keywords per domain class
confidence_normalization	max_keywords	["max_keywords," "sum_keywords"]	Confidence score normalization

Table 8. Hyperparameters of term frequency–inverse document frequency

Parameter	Default value	Range/options	Description
max_features	5,000	1,000–10,000	Maximum vocabulary size
ngram_range	(1, 2)	[(1,1), (1,2), (2,2)]	N-gram range
min_df	2	1–5	Minimum document frequency
max_df	0.8	0.7–1.0	Maximum document frequency
stop_words	English	["English," none, custom]	Stop-word list

lower variance compared to the baseline model, indicating enhanced predictive stability and robustness, as illustrated in **Figure 3**. In addition, the tighter distribution and higher median observed in the cross-validation results highlight improved predictive consistency across validation folds, as shown in **Figure 4**. Furthermore, the neuro-symbolic model achieved a substantially lower ECE (0.0444) compared to the baseline model (ECE = 0.1200), representing a 63% reduc-

Table 9. Logistic regression hyperparameters

Parameter	Default value	Range/options	Description
C	1.0	0.001–1,000	Inverse regularization strength
solver	liblinear	[“liblinear,” “lbfgs,” “saga”]	Optimization algorithm
max_iter	1,000	100–5,000	Maximum iterations
class_weight	balanced	[“balanced,” none]	Class weight handling
penalty	l2	[“l1,” “l2,” “elasticnet”]	Regularization type

Table 10. Neural network hyperparameters

Parameter	Default value	Range/options	Description
vocab_size	5,000	1,000–20,000	Vocabulary size for embedding
embedding_dim	100	50–300	Word embedding dimension
max_sequence_length	100	50–500	Maximum text sequence length
lstm_units	64	32–256	LSTM hidden units
bidirectional	True	True/false	Use bidirectional LSTM
dropout_rate	0.5	0.0–0.7	Dropout rate for regularization
recurrent_dropout	0.2	0.0–0.5	Recurrent dropout rate
dense_units_1	128	64–512	First dense layer units
dense_units_2	64	32–256	Second dense layer units
activation	relu	[“relu,” “tanh,” “sigmoid”]	Activation function
learning_rate	0.001	0.0001–0.01	Optimizer learning rate
batch_size	32	16–128	Training batch size
epochs	30	10–100	Training epochs
early_stopping_patience	5	3–10	Early stopping patience

Abbreviation: LSTM: Long short-term memory.

Table 11. Neuro-symbolic fusion hyperparameters

Parameter	Default value	Range/options	Description
feature_fusion	concatenation	[“concatenation,” “weighted,” “attention”]	Feature fusion method
symbolic_weight	0.5	0.0–1.0	Weight for symbolic features
neural_weight	0.5	0.0–1.0	Weight for neural features
confidence_threshold	0.7	0.5–0.9	Confidence threshold for decisions

Table 12. Evaluation hyperparameters

Parameter	Default value	Range/options	Description
test_size	0.2	0.1–0.3	Test set proportion
val_size	0.1	0.05–0.2	Validation set proportion
cv_folds	5	3–10	Cross-validation folds
random_state	42	Any integer	Random seed for reproducibility
scoring_metric	f1_macro	[“accuracy,” “f1_macro,” “roc_auc”]	Primary evaluation metric

tion in calibration error. This improvement indicates significantly better calibration of predicted probabilities and greater reliability in model confidence estimates, as presented in **Figure 5**. The learning curve analysis revealed that the neuro-symbolic approach achieved faster convergence,

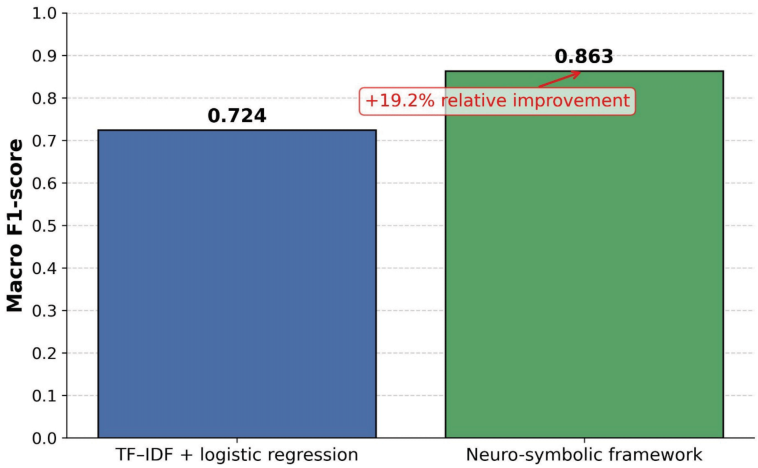

Figure 3. Comparison of macro F1-score between the baseline and the proposed neuro-symbolic framework Abbreviation: TF–IDF: Term frequency–inverse document frequency.

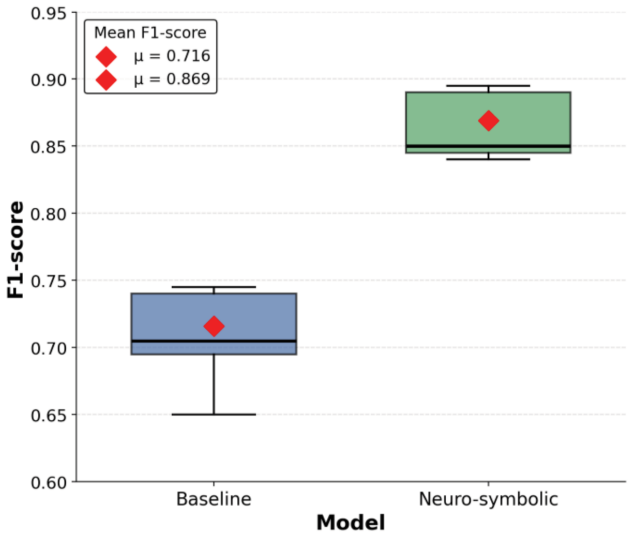

Figure 4. Ten-fold cross-validation performance comparison between the baseline and the proposed neuro-symbolic model

improved generalization, and greater stability as the size of the training dataset increased, as illustrated in **Figure 6**. These results collectively demonstrate the effectiveness of integrating neural learning with symbolic reasoning for robust and reliable biomedical text classification.

The results of this study illustrate a new generalization perspective: symbolic knowledge provides robust out-of-distribution performance while neural components capture pattern variation within known limits. In addition, this study contributes to the theoretical underpinning of calibrated AI systems by demonstrating that explicit reasoning capabilities improve not only accuracy but also the reliability of AI system estimates, which is essential to achieving trustworthy AI deployments in mission-critical environments.

To ensure methodological rigor and reproducibility, the proposed neuro-symbolic framework was evaluated using a comprehensive experimental protocol. The dataset of 7,569 full-length biomedical research articles was partitioned using a stratified split into 70% training, 15% validation, and 15% testing sets, preserving the distribution of the three classes (Colon_Cancer, Lung_Cancer, and Thyroid_Cancer). In addition, model performance was further validated using 10-fold stratified cross-validation, with reported results representing the average across all folds to ensure robust estimates of generalization. Hyperparameters were optimized via grid search on the validation set, including regularization and n-gram parameters for the TF–IDF logistic regression baseline, as well as the neural-symbolic fusion weight for the proposed framework. Model calibration was analyzed using ECE and reliability diagrams to evaluate the consistency between predicted probabilities and observed accu-

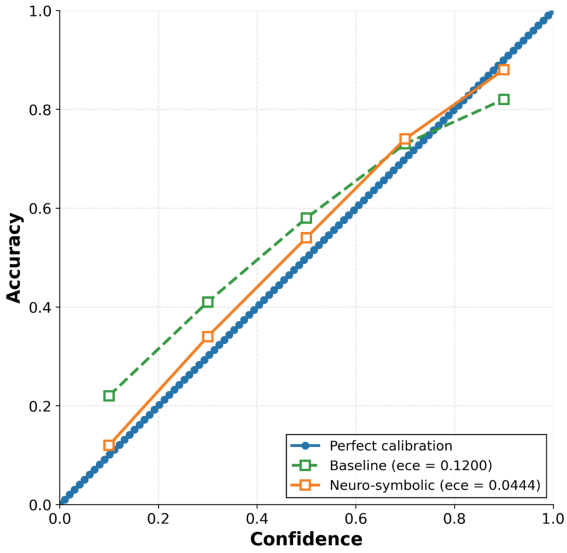

Figure 5. Reliability diagram comparing calibration performance of the baseline and neuro-symbolic models

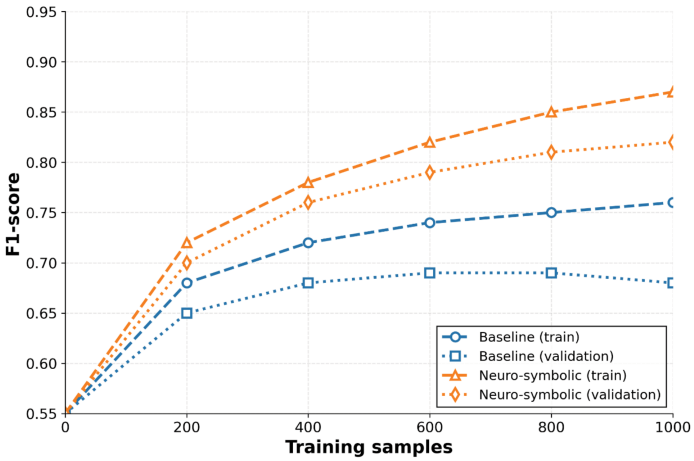

Figure 6. Learning curves comparison between the baseline and neuro-symbolic models

racy. Statistical significance of performance differences was assessed using a paired t-test across the cross-validation folds, with significance assessed at $\alpha = 0.05$ and effect size measured using Cohen's d.

The improvement in calibration enables better risk stratification and resource allocation, and the explicit reason-giving capabilities may increase clinicians' trust and ultimately acceptance of the technology. Additionally, the neuro-symbolic framework may be used in medical literature mining, where it can extract statistically associated patterns and provide them in an ontologically consistent representation, and in developing individualized treatment plans by encoding patient-specific constraints into symbolic representations and using the neural portion of the framework to probabilistically predict treatment outcome. The framework's ability to continually integrate new medical knowledge without completely retraining the model offered a necessary solution to the ongoing challenge of creating adaptable healthcare systems that evolve with advances in medical science.

5. Conclusion

The results demonstrate that neuro-symbolic hybridization can enhance the performance, interpretability, and reliability of AI systems applied to biomedical text analysis. By integrating pattern-based neural models with symbolic reasoning mechanisms, the proposed framework addressed several limitations of purely data-driven approaches, including improved logical consistency, more transparent decision explanations, and enhanced robustness in the presence of limited or imbalanced data. Experimental results show a 19.2% improvement in macro-F1 score, improved probability calibration, and stronger performance in detecting underrepresented classes.

These findings highlight the potential benefits of combining statistical learning with structured domain knowledge for biomedical document classification tasks. However, the experiments conducted in this study focused specifically on biomedical research literature; therefore, the results should be interpreted as methodological validation rather than direct clinical deployment. Future work should evaluate the proposed framework on clinical text sources, such as electronic health records or clinical narratives, to further assess its applicability in healthcare environments. In addition, challenges such as the computational cost of hybrid models, dependence on the completeness of the knowledge base, and rule conflict management remain important areas for further research. Future studies may explore automated knowledge acquisition, adaptive rule management, and improved uncertainty handling between neural and symbolic components to enhance the scalability and robustness of neuro-symbolic biomedical NLP systems.

Acknowledgments

The authors extend their sincere appreciation to Prince Sattam bin Abdulaziz University for funding this research through project number (PSAU/2025/01/36818).

Funding

None.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions

Conceptualization: All authors
Formal analysis: All authors
Methodology: All authors
Writing–original draft: All authors
Writing–review & editing: All authors

Availability of data

All data used in this study are publicly available from: https://www.kaggle.com/datasets/iamtanmayshukla/medical-text-classification-using-nlp/data

AI tools statement

All authors confirm that no AI tools were used in the preparation of this manuscript.

References

1. Bhuyan BP. Neuro-symbolic knowledge hypergraphs: knowledge representation and learning in neuro-symbolic artificial intelligence. Doctoral dissertation. Université Paris-Saclay; University of Petroleum and Energy Studies; 2025. Accessed April 11, 2026. https://theses.hal.science/tel-05051386/
2. Isakov A, Zaglubotskii A, Tomilov I, et al. Bridging heterogeneous agents: a neuro-symbolic knowledge transfer approach. *Technologies*. 2025;13(12):568. https://www.doi.org/10.3390/technologies13120568
3. Alqutaibi AY, AL-Zaghruri AS. The rise of AI in healthcare: are chatbots ready to lead? *Clin eHealth*. 2025;8:175-176. https://www.doi.org/10.1016/j.ceh.2025.09.001
4. García-Barragán Á, Sakor A, Vidal ME, et al. NSSC: a neuro-symbolic AI system for enhancing accuracy of named entity recognition and linking from oncologic clinical notes. *Med Biol Eng Comput*. 2025;63(3):749-772. https://www.doi.org/10.1007/s11517-024-03227-4
5. Shalaby AM, Hassan SMA, Abdelnour HM, et al. Ameliorative potential of bone marrow-derived mesenchymal stem cells versus prednisolone in a rat model of lung fibrosis: a histological, immunohistochemical, and biochemical study. *Microsc Microanal*. 2024;30(3):539-551. https://www.doi.org/10.1093/mam/ozae043
6. Azmat H, Maheen A. Neuro-symbolic approaches in NLP: integrating logic and learning for transparent language understanding. *J Integr Res*. 2025;6(3):1-9. Accessed April 11, 2026. https://interresearcher.com/index.php/JIR/article/view/3
7. Alghauli MA, Aljohani W, Almutairi S, Aljohani R, Alqutaibi AY. Advancements in digital data acquisition and CAD technology in dentistry: innovation, clinical impact, and promising integration of artificial intelligence. *Clin eHealth*. 2025;8:32-52. https://www.doi.org/10.1016/j.ceh.2025.03.001
8. Jammal F, Dahab M, Bayahya AY. Neuro-Bridge-X: a neuro-symbolic vision transformer with Meta-XAI for interpretable leukemia diagnosis from peripheral blood smears. *Diagnostics*. 2025;15(16):2040. https://www.doi.org/10.3390/diagnostics15162040
9. Singh S. Neuro-fuzzy architectures for interpretable AI: a comprehensive survey and research

outlook. *J Mach Learn Res*. 2025;1:11. Accessed April 11, 2026. https://www.preprints.org/frontend/manuscript/6176c0df3fe6e770725b0318288a0e8a/download_pub

10. Anand S, Miglani S, Anand R. Neuro-symbolic signal processing: a modular framework for adaptive and transparent real-time cognitive signal interpretation. *Int J Comput Appl*. 2025;186(66):24-30. https://www.doi.org/10.5120/ijca2025924351
11. Confalonieri R, Guizzardi G. On the multiple roles of ontologies in explanations for neuro-symbolic AI. *Neurosymbolic Artif Intell*. 2025;1:NAI-240754. https://www.doi.org/10.3233/NAI-240754
12. Alikhani MH. Synthetic reasoning—designing AI architectures beyond neural networks with hybrid neuro-symbolic systems. *SSRN*. 2025. https://www.doi.org/10.2139/ssrn.5226493
13. Anani M, Amer SA, Kishk RM, et al. Evaluation of blood and biochemical parameters of COVID-19 patients in Suez Canal University Hospital: a retrospective study. *J Infect Dev Ctries*. 2022;16(4):592-599. https://www.doi.org/10.3855/jidc.14591
14. Cheng K, Ahmed NK, Rossi RA, Willke T, Sun Y. Neural-symbolic methods for knowledge graph reasoning: a survey. *ACM Trans Knowl Discov Data*. 2025;18(9):1-44. https://www.doi.org/10.1145/3686806
15. Gao L, Sun L, Sheng VS. Enhancing transcription factor prediction via domain knowledge integration with logic tensor networks. *IEEE Trans Comput Biol Bioinform*. 2025;22(6):3564-3570. https://www.doi.org/10.1109/TCBBIO.2025.3617864
16. De Boer M, Smit Q, van Bekkum M, Meyer-Vitali A, Schmid T. Design patterns for large language model-based neuro-symbolic systems. *Neurosymbolic Artif Intell*. 2025;1. https://www.doi.org/10.1177/29498732251377499
17. Sabou M, Llugiqi M, Ekaputra FJ, Waltersdorfer L, Tsaneva S. Knowledge engineering in the age of neurosymbolic systems. *Neurosymbolic Artif Intell*. 2025;1. https://www.doi.org/10.1177/29498732251320078
18. Bhatnagar R, Hassan E, Saber A, Shams MY, Elbedwehy S. Enhanced vision transformer with Lion optimizer for accurate classification of monkeypox skin images. In: *Proc 4th Int Conf Technol Advancements Comput Sci (ICTACS)*. IEEE; 2024:1897-1902. https://www.doi.org/10.1109/ICTACS62700.2024.10840567
19. Bougzime O, Cruz C, André JC, Zhou K, Qi HJ, Demoly F. Neuro-symbolic artificial intelligence in accelerated design for 4D printing: status, challenges, and perspectives. *Mater Des*. 2025;252:113737. https://www.doi.org/10.1016/j.matdes.2025.113737
20. Zhang J, Chen B, Zhang L, Ke X, Ding H. Neural, symbolic and neural-symbolic reasoning on knowledge graphs. *AI Open*. 2021;2:14-35. https://www.doi.org/10.1016/j.aiopen.2021.03.001
21. Hassan E, Ghazalah SA, El-Rashidy N, et al. DenseNet model with attention mechanisms for robust date fruit image classification. *Int J Comput Intell Syst*. 2025;18(1):228. https://www.doi.org/10.1007/s44196-025-00809-4
22. Wu M, Li M, Qian C, et al. Enhance deep learning in few-shot datasets: the role of implicit symbolic rules. *Signal Image Video Process*. 2025;19(12):1013. https://www.doi.org/10.1007/s11760-025-04603-9
23. Mundlamuri R, Gunnam GR, Mysari NK, Pujuri J. The evolution of AI: from classical machine learning to modern large language models. *IEEE Access*. 2025;13:178302-178341. https://www.doi.org/10.1109/ACCESS.2025.3621344
24. Fathy MA, Anbaig A, Aljafil R, et al. Effect of liraglutide on osteoporosis in a rat model of type 2 diabetes mellitus: a histological, immunohistochemical, and biochemical study. *Microsc Microanal*. 2023;29(6):2053-2067. https://www.doi.org/10.1093/micmic/ozad102
25. Hassan E, El-Rashidy N, Elbedwehy S, et al. Exploring the frontiers of image super-resolution: a review of modern techniques and emerging applications. *Neural Comput Appl*. 2025;37(22):17913-17961. https://www.doi.org/10.1007/s00521-025-11331-1
26. Shams MY, Hassan E, Gamil S, et al. Skin disease classification: a comparison of ResNet50, MobileNet, and Efficient-B0. *J Curr Multidiscip Res*. 2025;1(1):1-7. https://www.doi.org/10.21608/jcmr.2025.327880.1002
27. Shalaby AM, Shalaby RH, Alabiad MA, et al. Evening primrose oil attenuates oxidative stress, inflammation, fibrosis, apoptosis, and ultrastructural alterations induced by metanil yellow in rat liver. *Ultrastruct Pathol*. 2023;47(3):188-204. https://www.doi.org/10.1080/01913123.2023.2189987
28. Alam F, Malik K, Krishnamurthy M. NeuroClustr: empowering biomedical text clustering with neuro-symbolic intelligence. In: *Proc IJCAI Workshop Knowledge-Based Compositional Generalization*. 2023. Accessed April 11, 2026. https://openreview.net/pdf?id=eu12MV19ho
29. Rajalakshmi R, Unhelkar B, Shankar S. Neuro-symbolic integration using knowledge attention graphs with advanced deep learning techniques for detecting brain disorders. *Int Insurance Law Rev*. 2025;33(S5):420-436.

https://www.doi.org/10.65677/iilr.33.S5.27
30. Hassan E, Bhatnagar R, Shams MY. Advancing scientific research in computer science by ChatGPT and LLaMA—a review. In: *Int Conf Intelligent Manufacturing and Energy Sustainability*. Springer; 2023:23-37. https://www.doi.org/10.1007/978-981-99-6774-2_3
31. Lu Q, Li R, Sagheb E, et al. Explainable diagnosis prediction through neuro-symbolic integration. *AMIA Jt Summits Transl Sci Proc*. 2025:332-341. https://www.doi.org/10.48550/arXiv.2410.01855
32. Keber M, Grubišić I, Barešić A, Jović A. A review on neuro-symbolic AI improvements to natural language processing. In: *2024 47th MIPRO ICT and Electronics Convention (MIPRO)*. IEEE; 2024:66-72. https://www.doi.org/10.1109/MIPRO60963.2024.10569741
33. Jain M, Singh K, Mutharaju R. ReOnto: a neuro-symbolic approach for biomedical relation extraction. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer; 2023;14172:230-247. https://www.doi.org/10.1007/978-3-031-43421-1_14
34. Alshahrani M, Khan MA, Maddouri O, et al. Neuro-symbolic representation learning on biological knowledge graphs. *Bioinformatics*. 2017;33(17):2723-2730. https://www.doi.org/10.1093/bioinformatics/btx275
35. Lucos R, LP S, PR V, et al. Neuro-symbolic artificial intelligence: foundations, advances, and future directions. 2025. https://www.doi.org/10.2139/ssrn.5915242
36. Jung YJ, Subramani P, Ashokkumar S. Neuro-symbolic cognition for intelligent autonomy and adaptive interpretability in wearable healthcare systems. *IEEE Trans Consum Electron*. 2025. https://www.doi.org/10.1109/TCE.2025.3647696
37. Prenosil GA, Weitzel TK, Bello SC, et al. Neuro-symbolic AI for auditable cognitive information extraction from medical reports. *Commun Med*. 2025;5(1):491. https://www.doi.org/10.1038/s43856-025-01194-x
38. Gueddes A, Fathallah W, Mahjoub MA. OntoMed-KGTransformer: a neuro-symbolic framework for clinical knowledge fusion. *Procedia Comput Sci*. 2025;270:426-435. https://www.doi.org/10.1016/j.procs.2025.09.161
39. Mileo A. Towards a neuro-symbolic cycle for human-centered explainability. *Neurosymbolic Artif Intell*. 2025;1:NAI-240740. https://www.doi.org/10.3233/NAI-240740
40. Bhuyan BP, Ramdane-Cherif A, Singh TP, Tomar R. Neuro-symbolic AI in various domains. In: *Neuro-Symbolic Artificial Intelligence: Bridging Logic and Learning*. Springer; 2024:311-324. https://www.doi.org/10.1007/978-981-97-8171-3_17

Zahraa Tarek *received the Ph.D. in computer science from Mansoura University, Egypt. She is currently a professor in information technology with the Faculty of Computers and Artificial intelligence, Mansoura University, Egypt. She has published many research articles in prestigious international conferences and reputable journals. She is also a reviewer for many journals. Her ongoing work continues to support both academic innovation and practical technological advancements across the AI and Big Data landscape.*
https://orcid.org/0000-0001-9389-2850

Esraa Hassan *was born in Kafrelsheikh, Kafrelsheikh, Egypt, in 1992. She received a B.Sc. Degree in Computer Science and a M.Sc. Degree in Computer Science from Mansoura University, Egypt, in 2013 and 2018, respectively. She received her Ph.D. in computer science from Mansoura University, Egypt, in 2022. She is currently a lecturer at the Faculty of Artificial Intelligence, Kafrelsheikh University, Egypt. She has published many research articles in prestigious international conferences and reputable journals. She is also a reviewer for many journals. Her research interests include machine learning, deep learning, optimization, and bioinformatics.*
https://orcid.org/0000-0002-1021-717X