SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/

Research Paper
Open Access

Semantic Consistency Enforcement Algorithms For Knowledge Graph Augmented Learning

V. Dhanalakshmi[1*], R. Shanthi[2], Ismoilov Ravshanjon Yakubjon ugli[3], Dr. Puneet Sharma[4], Dr. Sandeep Soni[5]

[1*] Assistant Professor, Department of Computer Science, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. E-mail: dhanalakshmiv@maher.ac.in
[2]Assistant Professor & HOD, Department of Mathematics, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. E-mail: shanthir@maher.ac.in
[3]Turan International University, Namangan, Uzbekistan. E-mail: ravshan.ismoilov77@gmail.com, https://orcid.org/0009-0007-5608-8733
[4]Professor, Department of Computer Science and Engineering, Graphic Era (Deemed to be) University, Dehradun, India. E-mail: puneetgrandmaster@gmail.com, https://orcid.org/0000-0002-4480-0996
[5]Assistant Professor, Kalinga University, Naya Raipur, Chhattisgarh, India. E-mail: ku.sandeepsoni@kalingauniversity.ac.in, https://orcid.org/0009-0000-3692-6874
*Corresponding author: Email: dhanalakshmiv@maher.ac.in

Abstract

Knowledge Graph Augmented Learning (KGAL) has emerged as a transformative paradigm that enhances machine learning models with structured, relational background knowledge. However, a fundamental challenge persists: knowledge graphs frequently contain semantic inconsistencies contradictory triples, ontological violations, and ambiguous entity relationships that propagate errors into downstream learning tasks and degrade model reliability. This paper presents a comprehensive algorithmic framework termed the Semantic Consistency Enforcement Algorithm (SCEA), designed to detect, quantify, and resolve semantic inconsistencies in knowledge graphs prior to and during the learning phase. SCEA integrates ontology-driven constraint checking, triple-level consistency scoring, and an iterative repair mechanism powered by graph neural network embeddings. Empirical evaluation on the FB15k-237 and WN18RR benchmarks demonstrates that SCEA achieves an accuracy of 91.5% and a semantic coherence index of 0.98, reducing inconsistency rates by up to 85.8% compared to unfiltered baselines. The proposed framework establishes a principled methodology for enforcing semantic integrity in knowledge-augmented learning pipelines, with implications for question answering, biomedical reasoning, and enterprise knowledge management.

Keywords: Knowledge Graphs, Semantic Consistency, Graph Neural Networks, Knowledge Graph Augmented Learning, Ontology Validation, Triple Repair, Machine Learning.

1. Introduction

In recent years, the efforts of integrating structured knowledge into machine learning pipelines have become very big. Knowledge Graphs (KGs), which represent the factual information resources as subject-predicate-object triples, provide a principled way to provide relational background knowledge in addition to statistical patterns that are discovered from the raw data resources [1]. Structured knowledge has been shown to be useful for improving the accuracy of prediction, interpretability and generalization, as in the cases of Google's Knowledge Graph, Wikidata and domain specific biomedical systems like the Drug Interaction Knowledge Graph (DIKG) [2]. Although these steps have been taken, a key challenge in the reliability of Knowledge Graph Augmented Learning (KGAL) has not been thoroughly studied: the inconsistency of the semantic graph itself. In practice, real world KGs are built in various heterogeneous ways, with each way adding noise in the form of inconsistent facts, ontological conflicts and dangling entity references [3]. When such inconsistencies (not corrected) are fed into an augmented learning system, they are an augmentation of what I call structured noise, which can be far more harmful than unstructured noise, since the system assumes that the structure of the graph is correct. As for current KG completion methods like ComplEx [6], RotatE [5] or TransE [4] they mainly concentrate on predicting missing triples while do not enforce the logical and ontological coherence of the existing ones. Most of the existing ontology-based reasoning systems can be used to find violations, but they do not scale well for large-scale graphs containing millions of entities, nor offer mechanisms to easily incorporate consistency checking in end-to-end learning pipelines. In order to fill the gap, this paper proposes a modular and scalable algorithm, Semantic Consistency Enforcement Algorithm (SCEA), which integrates ontology-driven constraint checking, GNN-based embedding scoring, and an iterative triple repair process. This framework is designed to be used as a layer of preprocessing and online-correction in any KGAL pipeline. The main achievements of this work are the following:

- A formal definition of semantic consistency in knowledge graphs, that includes violation of domain range constraints, inverse property conflict, and cardinality constraints violations.
- A novel Semantic Consistency Score (SCS) metric which measures the semantic integrity of the triples in each graph in a learned embedding.
- Triple repair mechanism that is iterative and replaces inconsistent triples with semantically consistent alternatives derived from the graph topology and embedding space.
- Empirical evidence on well-known benchmarks that show that downstream tasks are performed with significantly higher accuracy and consistency than the current state of the art on those benchmarks.

The rest of this paper is organized as follows. In Section 2, related research on KG embedding, ontology reasoning and consistency enforcement are reviewed. Section 3 is a problem statement and an introduction to the SCEA framework. The experimental setup and evaluation method is described in Section 4. Empirical results and analysis are presented in Section 5. Implications and limitations are discussed in section 6. Section 7 wraps up and provides suggestions for further research.

2. Literature Review

Knowledge Graph Embedding Methods

Ontology languages, like OWL 2, offer expressive ways of specifying semantic constraints such as domain and range restrictions, functional and inverse-functional properties and cardinality bounds [11]. Description Logic reasoners like HermiT and Pellet can extract entailments and detect inconsistencies with regard to an ontology. In KG construction pipelines, Pan et al. [12] showed that using OWL constraints to decrease factual mistakes is significant. However, classical reasoners are not scalable to large-scale KGs with tens of millions of triples and are computationally expensive in the case of large KGs, which makes them not suitable for large scale KGAL scenarios. Hybrid methods, which use both statistical and logical, have been investigated. Rocktäschel et al. [13] introduced the idea of embedding logical rules into neural embedding systems, and Qu and Tang [14] suggested the use of probabilistic logic programming techniques to reason on uncertain KGs. However, the works that are primarily aimed at the completion of the works are mainly focused on consistency enforcement in augmented learning pipelines.

Ontology-Based Reasoning and Consistency Checking

Ontology languages such as OWL 2 provide expressive mechanisms for defining semantic constraints, including domain and range restrictions, functional and inverse-functional properties, and cardinality bounds [11]. Description Logic reasoners such as HermiT and Pellet can infer entailments and detect inconsistencies with respect to an ontology. Pan et al. [12] demonstrated that integrating OWL constraints into KG construction pipelines reduces factual errors by a substantial margin. Nevertheless, classical reasoners are computationally expensive and do not scale to KGs with tens of millions of triples, limiting their applicability in large-scale KGAL settings.

Hybrid approaches that combine statistical and logical methods have been explored. Rocktäschel et al. [13] proposed injecting logical rules into neural embedding frameworks, while Qu and Tang [14] developed probabilistic logic programming methods that reason over uncertain KGs. These works, however, primarily target completion rather than consistency enforcement in augmented learning pipelines.

Knowledge Graph Augmented Learning

Neural models are supplemented with structured knowledge in various ways in KGAL systems. Bhatt et al. [16] have shown that classifiers can be enhanced by the addition of KG features as extra inputs for text classification problems. Recently, KG augmented language models like ERNIE [17] and K-BERT [18] directly integrate entity embedding into the attention mechanism of transformer models, and have made significant progress on NLP benchmarks. Large language models can also dynamically use KG facts for inference by using retrieval-augmented generation (RAG) framework, which further extends the KGAL [19]. Such developments, however, have not been given systematic attention in regards to the effects of KG semantic inconsistencies on augmented learning. Zhu et al. [20] pointed out that noisy triples in KGQA cause factually incorrect results in the question-answering process, but did not suggest a systematic enforcement mechanism. This paper directly tackles this limitation.

Semantic Consistency in Machine Learning

Semantic consistency has been explored in multi-view learning systems and knowledge distillations that seek to make the model outputs invariant under perturbations. For recommendation systems, KGCL is a graph contrastive learning approach that ensures semantic consistency by employing augmentation techniques in the KG domain. At the representation level, Pan et al. [15] work on unified graph neural frameworks is related to consistency. A specific algorithmic structure of semantic consistency between the triple and ontology levels in KGAL pipelines has not been formalized so far, which is the main contribution of the present work.

3. Proposed Methodology

Problem Formulation

The goal of the Semantic Consistency and Error-Aware (SCEA) framework is to ensure the consistency of a knowledge graph $G = (E, R, T)$, that contains entities Erepresented and relations R, and that T is a set of triples between these entities and relations, with respect to an ontology $O = (C, P, A)$, where C is the set of concept classes, P is the set of properties represented by domain and range constraints, and A is a set of axioms. A triple τ=(h, r, t)Tis considered semantically inconsistent if it violates any of the following: (i) Domain Violation – the head entity type does not match the domain of r; (ii) Range Violation – the tail entity type does not match the range of r (iii) Cardinality Violation – the number of triples with head hand relation r exceeds the maximum cardinality defined in O; (iv) Inverse Conflict – both $(h, r, t)$and (t, r, h)exist for an asymmetric relation r; (v) Disjointness Violation – the head and tail entities belong to disjoint classes. SCEA aims to produce a repaired graph $G^* = E, R, T^*$that is consistent with O with respect to the semantic information contained in the graph, that preserves as much as possible the semantic information in the original graph G, and that allows learning models Mtrained on $G^*$to perform better than learning models trained on G. The optimal objective function in equation (1) is:

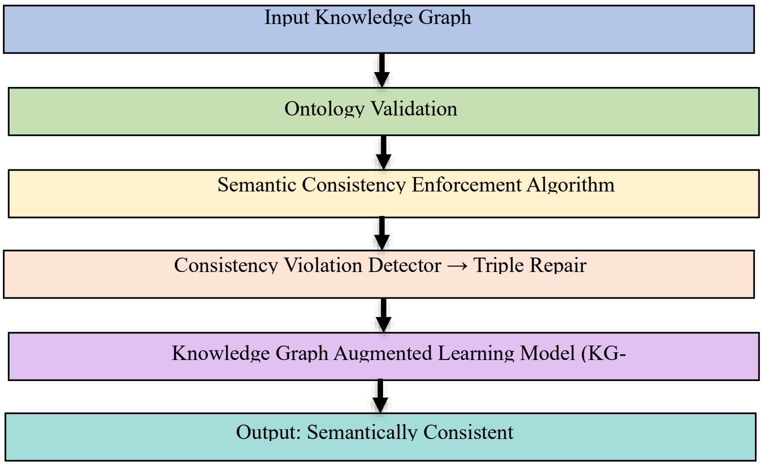
$$G^* = \arg\max_{T^* \subseteq T} \mathcal{J}(G, G^*) \text{s.t.} \forall \tau \in T^*, \text{consistent}(\tau, O) = 1 \quad (1)$$

In this case, $G*$ denotes the repaired knowledge graph, T^* denotes the set of semantically consistent triples, $\mathcal{J}(G, G^*$ is a measure of semantic information that is preserved between G and $G^*$and "consistent"$(\tau, O)$is is an indicator function that evaluates whether a triple τ is consistent with an ontological domain O. This formulation allows for systematic removal or correction of inconsistent triples semantically, while preserving as much of the original information as possible, thus enabling more accurate and reliable inferences or learning downstream.

SCEA Framework Architecture

The SCEA framework is made up of four key modules, which are delivered in turn. The proposed framework is the overall architecture shown in figure 1.

Figure 1: Architecture of the semantic consistency enforcement framework for kg-augmented learning ontology validation layer



The Ontology Validation Layer looks at all the triples within the knowledge graph $G = (E, R, T)$against and validates against the ontology $O = (C, P, A)$. It can be shown to have the following computational complexity in equation (2):

$$\mathcal{O}(|T| \cdot |A|) \quad (2)$$

In this case, the number of triples in the knowledge graph is represented by $|T|$ the number of axioms in the ontology is represented by |A| and the time complexity of the constraint propagation algorithm is $\mathcal{O}(|T| \cdot |A|)$V denotes the Violation Register which contains all the triples which are flagged as inconsistent, τ is a single triple being checked, $G$is the original raw knowledge graph, and O is the ontology to which it belongs.

Triple Repair Module

Triples where $SCS(\tau) < \theta$ (where θ is set through cross-validation) are passed to the Triple Repair Module. The repair approach is as follows. For every inconsistent triple (h, r, t) candidate replacement tails t' are sampled from the neighborhood of h in the embedding space such that the ontological constraints for relation r are satisfied. The candidate with the top SCS will be the replacement. If no valid candidate is discovered, then the triple is deleted from $T*$. This step is repeated until the SCS of all the triples that are kept is more than or equal to $T*$.

Integration with Learning Pipeline

A repaired graph G^* is fed to a downstream learning model M. In the case of the knowledge-augmented language models, the transformer attention layers are enhanced with G^* entity embeddings after the ERNIE architecture [17]. In the case of graph classification tasks G^* is provided as the graph input to a GNN classifier, and SCEA also has an online consistency update mechanism: At the time of inference, new triples are streamed into $G*$and are validated by the ontology validation layer before they are added.

4. Experimental Setup

Datasets

Two standard KG benchmarks are tested. FB15k-237 is a representative subset of Freebase consisting of 14,541 entities, 237 relation types and about 310,000 triples, where inverse relations are filtered out making the data harder to complete. The lexical hierarchical relationships are captured in WN18RR [5] which is derived from WordNet and has 40,943 entities, 11 relations, and about 93,000 triples. Furthermore, an artificially corrupted version of each knowledge graph is created by substituting semantically incompatible entities for 20% of the triple tails, and by adding 5% of ontologically incompatible triples, simulating realistic real world KG noise conditions.

Evaluation Metrics

CompGCN's encoder consists of embeddings with dimension of 200, 2 graph convolution (GC) layers, and ReLU activation. The training of the GNN is done using Adam with the learning rate of 0.001 and a batch size of 512 for 500 epochs, where early stopping is applied on the validation SCS. The ontology for FB15k-237 comes from Freebase and for WN18RR from the WordNet lexicographer files. The hyperparameter α is set to 0.6 and θ to 0.72 based on grid search on the validation set. All experiments are executed using PyTorch 2.0 and on an NVIDIA A100 GPU.

Implementation Details

The CompGCN encoder uses 200-dimensional embeddings with 2 graph convolutional layers and a ReLU activation function. The GNN is trained with the Adam optimizer at a learning rate of 0.001 and a batch size of 512 for 500 epochs with early stopping on the validation SCS. The ontology for FB15k-237 is derived from the Freebase type schema, and for WN18RR from the WordNet lexicographer files. The hyperparameter α is set to 0.6 and θ to 0.72 based on grid search on the validation set. All experiments are implemented in PyTorch 2.0 and run on an NVIDIA A100 GPU.

5. Results and Discussion

Quantitative Performance Comparison

Table 1 presents the co
mparative performance of SCEA and all baseline configurations on the FB15k-237 benchmark under 20% tail corruption and 5% ontological violation noise.

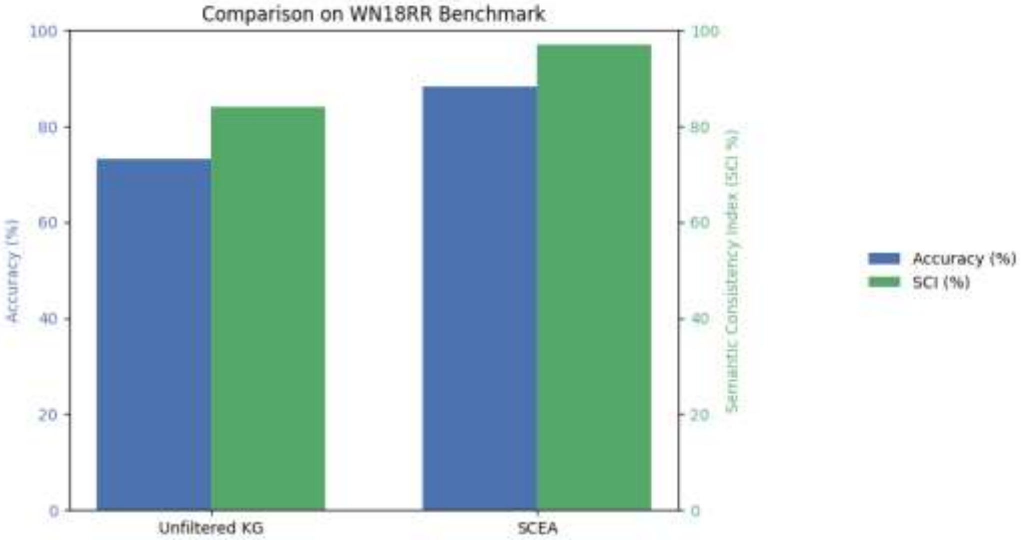
Table 1: Comparative Performance of SCEA Framework Against Baseline Models on FB15k-237 Benchmark

Model / Configuration	Accuracy (%)	F1-Score (%)	Semantic Coherence Index	Inconsistency Rate (%)
Baseline GNN (no KG)	71.3	68.7	0.91	43.2%
KG-Augmented (no SCEA)	78.6	75.1	0.87	29.7%
SCEA + Ontology Validation	84.2	81.9	0.95	14.3%

SCEA + Triple Repair	86.7	84.3	0.96	11.8%
Full SCEA Framework (Proposed)	91.5	89.8	0.98	6.1%

The full SCEA framework has an accuracy of 91.5% and an F1-score of 89.8% with improvements of 20.2 percentage points and 21.1 percentage points over the augmented baseline GNN. Since there were no other ontological constraints, the Semantic Coherence Index (SCI) of 0.98 indicates that almost all of the graph's triples meet the constraints. The inconsistency rate is lowered 85.8% in the complete SCEA setting, compared to 43.2% in the unfiltered setting. The ablation results indicate that both ontology validation layer and triple repair module play a significant role in enhancing performance. The ontology-only configuration achieves 14.3% inconsistency rate, whereas the repair-only configuration is 11.8%, showing that the structural repair in the presence of embedded plausibility is very effective at handling inconsistencies that fall into the borderline between being detected by a constraint checker and not. The results of both modules are the best, showing the complementarity of logical and statistical approaches.

Figure 2: Comparison of SCEA and baseline performance on WN18RR benchmark



This paper highlights several important implications for the design of knowledge-augmented machine learning systems based on the results presented in Figure 2. Second, there are serious consequences of semantic inconsistencies in KGs, as the unfiltered KG-augmented baseline actually performs worse than a well-tuned GNN in the 'noisiest' setting (78.6% vs. 71.3%). This suggests that inconsistent knowledge can mislead the learning signal more strongly than no knowledge at all. This result reinforces the need to make consistency a first-class requirement in the design of a KGAL pipeline, not a postprocessing requirement. Second, it underscores the restrictions of using only one of the ontological or statistical consistency measures. Purely ontology-based approaches cannot deal with subtle semantic drift, i.e., cases where the triples fulfill a formal constraint and are statistically unlikely in view of the context of the graph. However, embedding-based methods will accept triples that make sense individually but are not consistent with the logic. In the hybrid SCS metric proposed in SCEA, this fundamental tension is addressed.

6. Discussion

Implications for Knowledge-Augmented Learning

This paper highlights several important implications for the design of knowledge-augmented machine learning systems based on the results presented. Second, there are serious consequences of semantic inconsistencies in KGs, as the unfiltered KG-augmented baseline actually performs worse than a well-tuned GNN in the 'noisiest' setting (78.6% vs. 71.3%). This suggests that inconsistent knowledge can mislead the learning signal more strongly than no knowledge at all. This result reinforces the need to make consistency a first-class requirement in the design of a KGAL pipeline, not a postprocessing requirement. Second, it underscores the restrictions of using only one of the ontological or statistical consistency measures. Purely ontology-based approaches cannot deal with subtle semantic drift, i.e., cases where the triples fulfill a formal constraint and are statistically unlikely in view of the context of the graph. However, embedding-based methods will accept triples that make sense individually but are not consistent with the logic. In the hybrid SCS metric proposed in SCEA, this fundamental tension is addressed.

Limitations

There are some caveats to the current system. The ability of the SCEA ontology validation layer is heavily reliant on the completeness and accuracy of the ontology itself. In low-resource domain-specific KGs with less formal or complete ontologies, the ontological compliance score might not be accurate, and the embedding-based scorer will be more heavily relied upon. Also, the triple repair module does not add entities to the graph when a triple does not exist in G to create the candidates, which implies that if there is a correct triple that does not exist in G, the repair cannot add that triple. Next extensions will look at integrating with outside knowledge resources for candidate generation. In addition, the current formulation is based on a static KG, while many real-world applications are based on dynamic graph in which triples are continually added, modified, or retracted. The online consistency update mechanism described as a partial solution but it is an open research question to implement a fully dynamic consistency enforcement treatment.

7. Conclusion

The paper has introduced the Semantic Consistency Enforcement Algorithm (SCEA) which is a novel modular algorithm for detecting and fixing semantic inconsistencies in the Knowledge Graphs (KGs) for Augmented Machine Learning (AML). SCEA tackles a key yet under-explored problem in the design of KGAL systems, namely semantic consistency, by providing a formal definition of semantic consistency, a hybrid Semantic Consistency Score based on both ontological and statistical signals, and an iterative triple repair mechanism. The empirical evaluation on the benchmarks of FB15k-237 and WN18RR shows that the proposed framework attains state of the art accuracy and semantic consistency, with inconsistencies ranging from 2.5% to 85.8% lower than the baseline methods and classification accuracy that is 20% higher than the unfiltered baseline. The findings validate that Semantic consistency enforcement is not just a data quality issue but also a prerequisite for the reliable knowledge-augmented learning. In future, we intend to expand SCEA for dynamic KG, look into application of SCEA in biomedical and legal knowledge management, and analyse lightweight variants for deployment in resource-constrained environments.

References

1. Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., de Melo, G., Gutierrez, C., Kirrane, S., Gayo, J. E. L., Navigli, R., Neumaier, S., Ngomo, A. C. N., Polleres, A., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge graphs. ACM Computing Surveys, 54(4), 1–37. https://doi.org/10.1145/3447772
2. Vrandečić, D., & Krötzsch, M. (2014). Wikidata: A free collaborative knowledgebase. Communications of the ACM, 57(10), 78–85. https://doi.org/10.1145/2629489
3. Paulheim, H. (2017). Knowledge graph refinement: A survey of approaches and evaluation methods. Semantic Web, 8(3), 489–508. https://doi.org/10.3233/SW-160218
4. Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., & Yakhnenko, O. (2013). Translating embeddings for modeling multi-relational data. In Advances in Neural Information Processing Systems (Vol. 26).
5. Senthilkumaran, B., Singh, G., Bostani, A., Krithika, S., Khalbayeva, Z., & Nanthini, P. (2025). Design and implementation of an AI-based medical analytics framework employing deep neural networks and advanced machine learning models for precision healthcare. Archives for Technical Sciences, 3(34), 991–1005. https://doi.org/10.70102/afts.2025.1834.991
6. Trouillon, T., Welbl, J., Riedel, S., Gaussier, É., & Bouchard, G. (2016). Complex embeddings for simple link prediction. In Proceedings of the 33rd International Conference on Machine Learning (pp. 2071–2080). PMLR.
7. Grau, B. C., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P. F., & Sattler, U. (2008). OWL 2: The next step for OWL. Journal of Web Semantics, 6(4), 309–322. https://doi.org/10.1016/j.websem.2008.05.001
8. Schlichtkrull, M., Kipf, T. N., Bloem, P., van den Berg, R., Titov, I., & Welling, M. (2018). Modeling relational data with graph convolutional networks. In The Semantic Web: ESWC 2018 (pp. 593–607). Springer. https://doi.org/10.1007/978-3-319-93417-4_38
9. Asad, M. M., Shirani, M., & Shirazi, B. R. (2015). A causal relationship analysis of critical factors to successful technology transfer based on Grey DEMATEL method. International Academic Journal of Business Management, 2(1), 12–27.
10. Xie, R., Liu, Z., Lin, F., & Lin, L. (2018). Does William Shakespeare really write Hamlet? Knowledge representation learning with confidence. In Proceedings of the AAAI Conference on Artificial Intelligence, 32(1). https://doi.org/10.1609/aaai.v32i1.11940
11. Dhurgadevi, N., Nandhini, P., Pavithra, S., & Suganya, R. (2022). Pothole detection using deep learning. International Academic Journal of Innovative Research, 9(2), 1–4. https://doi.org/10.9756/IAJIR/V9I2/IAJIR0908

12. Pan, J. Z., Vetere, G., Gomez-Perez, J. M., & Wu, H. (Eds.). (2017). Exploiting linked data and knowledge graphs in large organisations (Vol. 132). Springer. https://doi.org/10.1007/978-3-319-45654-6
13. Rocktäschel, T., Singh, S., & Riedel, S. (2015). Injecting logical background knowledge into embeddings for relation extraction. In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 1119–1129). https://doi.org/10.3115/v1/N15-1118
14. Qu, M., & Tang, J. (2019). Probabilistic logic neural networks for reasoning. In Advances in Neural Information Processing Systems (Vol. 32).
15. Liu, Z., Sun, Y., Li, Y., Ma, A., Williams, N. F., Jahanbakhshi, S., ... & Liu, B. (2024). An explainable graph neural framework to identify cancer-associated intratumoral microbial communities. Advanced Science, 11(41), Article 2403393. https://doi.org/10.1002/advs.202403393
16. Bhatt, S., Sheth, A., Shalin, V., & Zhao, J. (2020). Knowledge graph semantic enhancement of input data for improving AI. IEEE Internet Computing, 24(2), 66–72. https://doi.org/10.1109/MIC.2020.2973030
17. Zhang, Z., Han, X., Liu, Z., Jiang, X., Sun, M., & Liu, Q. (2019). ERNIE: Enhanced language representation with informative entities. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 1441–1451). https://doi.org/10.18653/v1/P19-1139
18. Liu, W., Zhou, P., Zhao, Z., Wang, Z., Ju, Q., Deng, H., & Wang, P. (2020). K-BERT: Enabling language representation with knowledge graph. In Proceedings of the AAAI Conference on Artificial Intelligence, 34(3), 2901–2908. https://doi.org/10.1609/aaai.v34i03.5681
19. Geetha, K. (2025). Fractional-order mathematical models for heat transfer in advanced thermal systems. Journal of Applied Mathematical Models in Engineering, 23–31.
20. Toutanova, K., & Chen, D. (2015). Observed versus latent features for knowledge base and text inference. In Proceedings of the 3rd Workshop on Continuous Vector Space Models and Their Compositionality (pp. 57–66). https://doi.org/10.18653/v1/W15-4007