Proceedings of the 2023 International Conference on Machine Learning and Cybernetics, 9-11 July, 2023

A PERSPECTIVE SURVEY ON INDUSTRIAL KNOWLEDGE GRAPHS: RECENT ADVANCES, OPEN CHALLENGES, AND FUTURE DIRECTIONS

YUHU SHANG[1], YIMENG REN[1], HAO PENG[2], YUE WANG[2], GANG WANG[3], ZHONGCHENG LI[4], YANGZHAO YANG[5], YANGYANG LI[5]

[1]School of artificial Intelligence, Tianjin University of Science & Technology, Tianjin 300457,China
[2]School of Cyber Science and Technology, Beihang University, Beijing 1000191, China
[3]School of Automation, Beijing Institute of Technology, Beijing 100081, China
[4]China Tower Corporation Hebei Branch, China [5]Academy of Cyber, China
E-MAIL: shangyuhu6902@mail.tust.edu.cn, renyimeng@mail.tust.edu.cn, penghao@buaa.edu.cn, zb2039111@buaa.edu.cn
gangwang@bit.edu.cn, lizc@chinatowercom.cn, 47702413@qq.com, liyangyang@live.com

Abstract:

As a result of the development of a new generation of artificial intelligence and carbon-neutral technologies, traditional industries are undergoing dramatic transformations. The exploration of industrial intelligence is still in its nascent stages, particularly lacking technical approaches to distill experiential knowledge from heterogeneous data sources originating from various origins. Knowledge Graphs (KG), as cutting-edge artificial intelligence technologies, can enable knowledge management and reuse while condensing valuable knowledge. As a result, fully utilizing KG's potential in the industrial field is critical to the realization of autonomous sensing, cognition, and the evolution of next-generation intelligent manufacturing systems. This paper starts with an overview of the current state of industrial knowledge graph development and shows how to construct an industrial knowledge graph (IKG). Following that, we provide a thorough and in-depth review of various industrial scenarios supported by knowledge graphs. Furthermore, this paper identifies the current challenges confronting industrial applications and proposes future research directions for IKG. It is hoped that this research will draw the attention of more researchers to the knowledge graph-based smart manufacturing paradigm and benefit their work.

Keywords:

Artificial Intelligence; Intelligent Manufacturing Systems; Industrial Knowledge Graph

1. Introduction

Intelligent manufacturing technologies have been driven by the rapid development of artificial intelligence, industrial internet, edge computing, and other emerging information technologies. Countries all over the world have identified the development of a new generation of intelligent manufacturing technology as a critical technology for increasing global competitiveness. Intelligent manufacturing systems have initially possessed basic capabilities such as state monitoring and auxiliary decision-making following recent development. However, there are still flaws, such as an over-reliance on human intervention and difficulty completing complex tasks on one's own. This is because current intelligent manufacturing systems are incapable of summarizing and summarizing empirical knowledge in the same way that humans do.

Knowledge Graphs is an artificial intelligence technology that manages, coordinates, and distills all types of knowledge (e.g. mechanism, data, etc.) [1]. By organizing massive heterogeneous concepts with interconnected nodes and defining cross-level relationships with heterogeneous edges, knowledge graph can aid intelligent manufacturing systems in their ability to acquire and deduce detailed knowledge [2]. When constructing industrial knowledge graph (IKG), strategies such as multiple views are commonly used to define the attribute characteristics of the connected edges of heterogeneous nodes [3]. This type of processing becomes a powerful support for the downstream tasks of intelligent manufacturing systems. By providing support for industrial intelligence, knowledge graph technology not only promotes the development of the industrial sector, but also improves the quality of production and service.

Related Surveys and Our Contributions. In recent years, the critical review of KG applications in industrial scenarios has attracted surging attention. For instance, the authors of [4] particularly analyzed potential applications of KGs in scenarios and mainly focused on maintenance, optimization, and resource allocation. Studies related to KG for Industry 4.0 modeling standards, norms, and frameworks are reviewed in [5]. By contrast, several surveys focus on different types of industrial scenarios, such as food science and industry [6], power transformers [7], smart grid [8] and fault diagnosis [9]. The key aim of this survey is to give a overview of the recent development, challenges and future research directions of industrial knowledge graphs, fully demonstrating the practicability of KG-driven technology in the industrial smart service. The main aspects of this review are as follows:

- We first revisit the various steps involved in the industrial knowledge graphs construction process.
- We categorize the IKGs and present several major works and their corresponding models for each topic.
- We summarize some challenges and future research directions of adopting KGs in the industrial manufacturing process.

2. Building Industrial Knowledge Graphs

In the context of the formal model regarding the construction of an industrial knowledge graph can be categorized into three stages: knowledge construction, representation and reasoning, and using stages. The construction process is shown in Fig. 1.

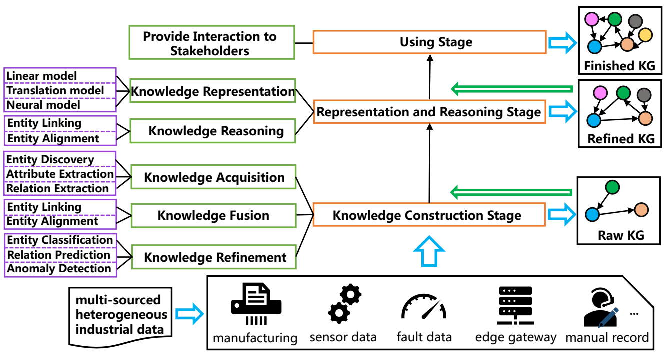

FIGURE 1. A workflow for constructing and implementing industrial knowledge graphs.

Knowledge Construction Stage. The raw data in industrial scenarios have a lot of noisy facts in the knowledge construction stage. Knowledge acquisition aims to acquire relevant entities, attributes, relations, rules, and facts from semi-structured and unstructured data in industrial scenarios. To guarantee the high accuracy of acquired knowledge, some advanced deep learning methods such as BiLSTM-CRF, BiLSTM-ALBERT, etc., which are also proposed to extract entities and indicate the dispatching behavior relationship patterns [10]. Knowledge fusion, achieved through entity alignment and entity linking, plays a crucial role in facilitating the sharing, association, and discovery of knowledge units [11]. Then, knowledge refinement can further detect noise in KGs to inconsistency checking and modify the knowledge graph adopting entity classification, relation prediction, and anomaly detection[12].

Representation and Reasoning Stage. Learning low-dimensional distributed embedding of entities and relations is the core of representation learning. In industrial scenarios, classic models for knowledge representation learning include translation-based [13] and tensor factorization-based [14] models, while neural network-based [15] methods have occupied an important position in this field. Then, the new information is obtained through knowledge reasoning[16], including entity alignment and padding and attribute value alignment.

Using Stage. When deploying knowledge systems into real-world applications, the primary concern of this stage is to enable stakeholders to manipulate the knowledge base without requiring detailed knowledge of its inner workings. Recently, numerous scholars have conducted on designing a human-machine interface that emphasizes ergonomic characteristics, such as friendliness and transparency [30]. For instance, chatbots [31], visualized graphs [32], and question-answering systems [33] have all proven to be effective in industrial cases.

3. An Overview of Knowledge Graphs Technology in Industrial Scenarios

Knowledge graphs contain rich structured knowledge and can be leveraged for various downstream tasks. The schematic diagram of industrial service scenario based on the industrial knowledge graph is shown in Fig. 2.

Predictive Maintenance. Predictive maintenance is a preventive maintenance approach that relies on online health assessment to predict the breakdown of a system by detecting early signs of failure. It is performed based on an online health assessment and predicts the breakdown of the system to be maintained by detecting early signs of failure. Nunez et al. [17] proposed an ontology-based predictive maintenance method, which utilizes sensor data, fault data, historical maintenance data, and other information to predict the future state of the machine and provide corresponding maintenance recommendations. Nunez et al. [18] also used ontology to define the relationship between machines and their health status, and used expert system rules to determine whether the machine is in a normal state, warning state, or failure state. Recently, deep learning algorithms have gained much attention in the field of predictive maintenance for rotary machinery systems. For instance, Hou et al.[19] used knowledge graph reasoning method and machine learning algorithm to construct a fault prediction model of elevator running system, providing references for elevator maintenance management. Aiming to provide a complete architecture which could be used in industrial IoT, Qiu et al. [20] extracted the features from the vibration signals and automatically identified the dynamic characteristics of the machine structure based on knowledge graph, so as to automatically to monitor the machine tool health condition.

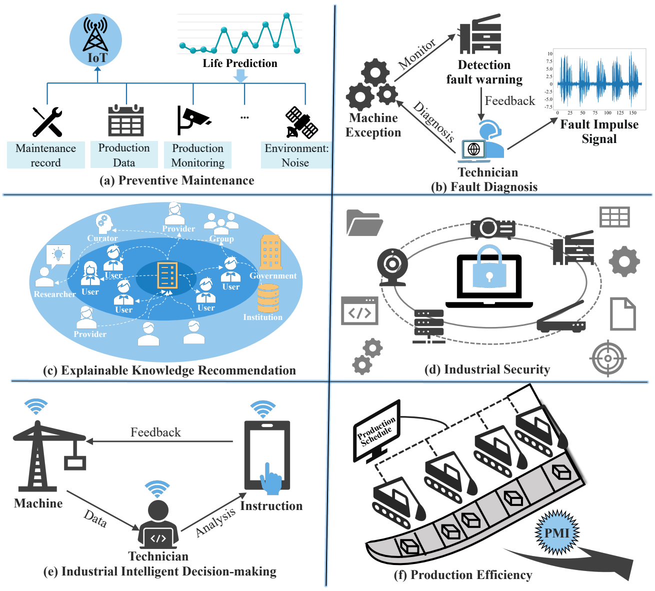

FIGURE 2. The schematic diagram of industrial service scenario based on the industrial knowledge graph.

Fault Diagnosis. Fault diagnosis plays a critical role in exploring the relationship between monitoring data and the health status of machines. For example, Hossayni et al. [21] applied distributed architecture and ontologies to diagnose and isolate faults in induction motors. Giustozzi et al. [22] presented an approach that uses knowledge graph stream reasoning to identify in real time certain situations that lead to potential faults. To solve the problem of diagnosing the root cause of abnormal performance in steam turbine applications, Qiu et al. [23] presented an root cause analysis method based on a knowledge graph and a bayesian network. Aiming at the problems of incomplete and uncertainty knowledge in machine tool fault diagnosis, Lu et al. [24] presented a data-driven iterative automatic construction method of knowledge graph, and applied fault knowledge graphs to assist fault troubleshooting applications. Wan et al. [26] developed a multi-faceted modelling approach that combines knowledge graph and machine learning algorithms to diagnose strip breakage during cold rolling in the steel industry. By fusing data from multiple sources, this approach provides a more comprehensive understanding of complex industrial processes and helps improve the accuracy of diagnosis. Given that manually collecting the labelling process from multiple sources corpus (such as maintenance logs and manuals) is time-consuming, Chen et al. [27] utilized a relation-aware-based sentence-level attention enhanced component for relation extraction. They trained the component on remote monitoring data as supervised learning samples to automatically construct a fault diagnosis knowledge graph. The core idea of the work [28] is innovative in combining causality mining and knowledge graph techniques to achieve root cause diagnosis of performance anomalies in cloud applications. Based on the Function-Behaviour-Structure theory, Shi et al. [29] presented an information integration approach for spacecraft fault diagnosis scenarios. Subsequently, deep reinforcement learning and graph neural network methods [25] are introduced, allowing the low-level graph to evolve into a high-level graph containing more abundant fault information.

Industrial Security. The large-scale deployment of sensors in industrial systems brings significant security risks and challenges. Dorodnykh et al. [35] proposed a method to automate the extraction of specific entities from tabular data and demonstrated the validity of the semantic interpretation of individual tabular elements as a key feature through a case study. Fang et al. [36] designed a hybrid computer vision and ontology model to identify hazards from images automatically. By using historical railway reports, Liu et al. [37] combined the data enhancement method to build a connected network of hazards and accidents to form a knowledge graph, which was applied to railway hazard identification and risk assessment. For numerical control machine tool fault analysis, Duan et al.[38] adopted risk-based knowledge graph to help production managers identify safety-critical components in CNC machine tools.

Intelligent Decision-making. In the last few years, intelligent decision-making has become a hot topic in industrial scenarios and benefited many decision-making-related tasks and real-world applications across a variety of areas. Munoz et al. [39] highlighted the potential of using knowledge graphs and multi-label learning models for adverse drug reaction prediction, which has important implications for improving drug safety and patient outcomes. Liu et al. [40] built multi-layer manufacturing domain knowledge graph based on the data collected from manufacturing process, thus achieving perception analysis and cognition decision-making in the resource allocation of the manufacturing process. The work in [41] introduced knowledge graph into an automatic machining process decision-making system, which is based on a three-level information model and a hybrid reasoning algorithm, leading to an effective system for intelligent process decision. In light of this, IKGs can assist organizations in making more informed decisions, reducing costs, and enhancing productivity in various industrial settings.

Production Efficiency. The increasing information complexity and its influence on production efficiency are fundamentally crucial for current manufacturing processes. Zhou et al. [46] explore the implicit relationship between complex engineering data through knowledge graph, it integrates the implicit engineering knowledge in a machining workshop environment and is used for supporting the optimization of resource allocation. Xu et al. [42] applied ontology and multiple decision graph (MDD) to cloud monitoring system to improve the running efficiency of the reasoning process.

4. Open Challenges

This section further presents some remaining challenges of adopting knowledge graphs on industrial scenarios.

Data Quality Requirements. The quality of data in practical industries has a significant impact on performance, and real industrial scenarios generally contain high levels of noise heterogeneous data [45]. In this case, development activities driven by noisy information may cause a system to crash or malfunction. To address these issues, data cleaning, normalization, and knowledge fusion techniques should be performed to remove inconsistencies and boost the quality of the data. Besides, some studies attempted to introduce other knowledge curation mechanisms to measure KG completion, and KG correctness approaches, but the involvement of extra work is still unavoidable to meet the infrastructure requirements.

Huge Computing Burden. The high complexity of industrial production requires large-scale IKGs, which in turn requires processing an enormous amount of training parameters. This leads to the problem of high computation cost, which is a pain point that plagues industrial applications. For instance, condition monitoring of robot gears can decrease unexpected downtimes of highly automated production lines, but the implementation of this solution is expensive due to costly hardware. Therefore, to better support low-cost industrial manufacturing, the interaction between knowledge graphs and the industry needs further enhancement. We are convinced that the development of a low-cost, high-efficiency industrial automation technology will be a great advance in this sector.

Multi-line and Multi-product Constraints. In industrial manufacturing production, a multitude of requirements and constraints must be considered, such as product line segments, resource sharing, minimum run-length constraints, and more, making it a highly complex decision-making problem. Existing methods usually adopt constraint optimization methods to optimize multiple objectives, reducing the loss of one objective when optimizing another. However, these methods optimize objective-by-objective, leading to time-consuming production processes. Therefore, industrial manufacturing production requires solutions that consider more constraints to execute many operations in industrial scheduling.

5. Future Perspectives

Based on the comprehensive literature survey above, we discuss several potential future research directions.

Human-machine Co-evolvement. Human workers excel at tasks requiring high-level cognitive decision-making and exception handling, while machines excel at precise computing, repeatability, and fast production. Using knowledge graphs, we can bridge the semantic gap and break human operators' thinking patterns with creative insights discovered by cognitive machines. In this sense, it is necessary to design more effective and elegant models to proactively improve their performance in the manufacturing process. Therefore, in the coming years, a sophisticated mechanism will be developed in which humans and machines can assist each other for sustainable growth.

Interpretable of Deep Model. Neural methods are scalable, but their lack of interpretability limits their application in industrial services, as operation managers may not fully trust models lacking explanatory results. The core idea of decision-making in engineering requires model outputs to be accompanied by scientific understanding. One straightforward way is utilizing interpretable methods, such as decision tree [47] and k-means clustering [48]. However, all these models are often less powerful than intelligent methods. In fact, knowledge graphs excel at processing multi-source information of industrial scenarios and providing interpretability. Further research on interpretability based on knowledge graphs is recommended to bridge the trust gap between operation managers and artificial intelligence.

Few-shot and Zero-shot Learning. The completeness and accuracy of IKGs largely depend on the quantity and quality of the data collected. However, in real-world industrial environments, the collected raw data generally have a skewed distribution. For instance, fault classification is an important issue in industrial monitoring, where one class, such as fault instances, is insufficiently represented compared to other classes, such as healthy instances. This results in relatively few fault instances in such datasets, which are likely classified as rare occurrences, ignored, or assumed as outliers, leading to inaccurate algorithm classification. Any manufacturing errors can cause irreparable damage to the industrial system. Therefore, there is a need to design more effective-based transfer learning models, such as knowledge distillation, to utilize industrial information better.

6. Conclusion

The industrial manufacturing scenario has a lot of complex industrial data, which has necessitated the development of advanced data analysis tools that are capable of handling the propagation and the heterogeneity of the above data. Therefore, we explored the domain of KG-based industrial and presented it well-categorized to express how a knowledge graph provides side information to industrial applications to enhance performance. In this survey, we introduced various methods proposed for building industrial knowledge graphs and discussed application-oriented approaches that use knowledge graphs to accomplish industrial tasks. Despite challenges in complex manufacturing scenarios, we have seen considerable application prospects for industrial knowledge graphs in the industrial domain. Lastly, we presented some future research perspectives to provide clear guidance for new researchers in this field.

Acknowledgements

This paper was supported by the National Key R&D Program of China through grant 2021YFB1714800, NSFC through grant 62002007, S&T Program of Hebei through grant 20310101D, Natural Science Foundation of Beijing Municipality through grant 4222030.

References

[1] M. Nickel, et al, “A review of relational machine learning for knowledge graphs”, Proc. IEEE, pp. 11-33, 2015.

[2] H. Liu, et al, “Machinery fault diagnosis based on deep learning for time series analysis and knowledge graphs”, Journal of Signal Processing Systems, Vol. 93, pp. 1433-1455, 2021.

[3] H. Ren, Z. Chen, et al, “An Industrial Multilevel Knowledge Graph-Based Local–Global Monitoring for Plant-Wide Processes”, TIM, Vol. 70, pp. 1-15, 2021.

[4] S. R. Bader, et al, “A knowledge graph for industry 4.0”, ESWC, pp. 465-480, 2020.

[5] N. Melluso, et al, “Enhancing Industry 4.0 standards interoperability via knowledge graphs with natural language processing”, Computers in Industry, 2022.

[6] A. Hassoun, S. Jagtap, H. Trollman, et al, “Food processing 4.0: Current and future developments spurred by the fourth industrial revolution”, Food Control, 2022.

[7] J. Martinez-Gil, G. Buchgeher, D. Gabauer, et al, “Root Cause Analysis in the Industrial Domain using Knowledge Graphs: A Case Study on Power Transformers”, Procedia Computer Science, pp. 944-953, 2022.

[8] H. Ding, et al, “A Review of the Construction and Application of Knowledge Graphs in Smart Grid”, iSPEC, pp. 3770-3775, 2021.

[9] W. Li, R. Huang, J. Li, Y. Liao, et al, “A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges”, Mechanical Systems and Signal Processing, 2022.

[10] B. Zhou, et al, “BA-IKG: BiLSTM Embedded ALBERT for Industrial Knowledge Graph Generation and Reuse”, INDIN, Vol. 1, pp. 63-69, 2020.

[11] D. Fisch, E. Kalkowski, et al, “Knowledge fusion for probabilistic generative classifiers with data mining applications”, TKDE, Vol. 26, No. 3, pp. 652-666, 2013.

[12] B. Zhou, J. Bao, Z. Chen, Y. Liu, B. Zhou, J. Bao, Z. Chen, and Y. Liu. KGAssembly: Knowledge graph-driven assembly process generation and evaluation for complex components. International Journal of Computer Integrated Manufacturing, 35(10-11):1151-1171, 2022.

[13] D. Jin, et al, “TransFusion: Multi-Modal Fusion for Video Tag Inference via Translation-based Knowledge Embedding”, ACM MM, pp. 1093-1101, 2021.

[14] A. B. M. Moniruzzaman, R. Nayak, M. Tang, et al, “Fine-grained type inference in knowledge graphs via probabilistic and tensor factorization methods”, WWW, pp. 3093-3100, 2019.

[15] Y. Lin, et al, “Dual contrastive prediction for incomplete multi-view representation learning”, TPAMI, 2022.

[16] B. Zhou, X. Shen, Y. Lu, X. Li, B. Hua, T. Liu, and J. Bao, “Semantic-aware event link reasoning over industrial knowledge graph embedding time series data”, International Journal of Production Research, pp.1-18, 2022.

[17] D. L. Nuñez, et al, “An ontology-based model for prognostics and health management of machines”, Journal of Industrial Information Integration, pp. 33-46, 2017.

[18] D. L. Nuñez, et al, “OntoProg: An ontology-based model for implementing Prognostics Health Management in mechanical machines”, Advanced Engineering Informatics, Vol. 38, pp. 746-759, 2018.

[19] J. Hou, et al, “Failure prediction of elevator running system based on knowledge graph”, Proceedings of the 3rd International Conference on Data Science and Information Technology, pp. 53-58, 2020.

[20] C. Qiu, et al, “A novel method for machine tool structure condition monitoring based on knowledge graph”, The International Journal of Advanced Manufacturing Technology, Vol. 120, pp. 563-582, 2022.

[21] H. Hossayni, et al, “SemKoRe: improving machine maintenance in industrial iot with semantic knowledge graphs”, Applied Sciences, Vol. 10, No. 18, 2020.

[22] F. Giustozzi, et al, “Abnormal situations interpretation in industry 4.0 using stream reasoning”, Procedia Computer Science, Vol. 159, pp. 620-629, 2019.

[23] N. Yang, et al, “Research on Knowledge Graph and Bayesian Network in Fault Diagnosis of Steam Turbine”, PHM-Shanghai, 2020.

[24] Y. Lu, et al, “Exploration of the construction and application of knowledge graph in equipment failure”, DEStech Transactions on Computer Science and Engineering, 2017.

[25] H. Peng, et al, “Spatial temporal incidence dynamic graph neural networks for traffic flow forecasting”, Information Sciences, Vol. 521, pp. 277-290, 2020.

[26] Y. Wan, et al, “Exploiting knowledge graph for multifaceted conceptual modelling using GCN”, Procedia Computer Science, Vol. 200, pp. 1174-1183, 2022.

[27] C. Chen, et al, “Reinforcement learning-based distant supervision relation extraction for fault diagnosis knowledge graph construction under industry 4.0”, Advanced Engineering Informatics, Vol. 55, 2023.

[28] J. Qiu, et al, “A causality mining and knowledge graph based method of root cause diagnosis for performance anomaly in cloud applications”, Applied Sciences, Vol. 10, No. 6, 2020.

[29] H. B. Shi, et al, “An information integration approach to spacecraft fault diagnosis”, Enterprise information systems, Vol. 15, No. 8, pp. 1128-1161, 2021.

[30] L. Ren, Y. Li, X. Wang, J. Cui, and L. Zhang. “An ABGE-aided manufacturing knowledge graph construction approach for heterogeneous IIoT data integration”, International Journal of Production Research, pp. 1-15, 2022.

[31] A. Sheth, H. Y. Yip, et al, “Extending patient-chatbot experience with internet-of-things and background knowledge: case studies with healthcare applications”, IEEE Intelligent Systems, Vol. 34, No. 4, pp. 24-30, 2019.

[32] J. Dou, J. Qin, et al, “Knowledge graph based on domain ontology and natural language processing technology for Chinese intangible cultural heritage”, Journal of Visual Languages & Computing, Vol. 48, pp. 19-28, 2018.

[33] Y. Zou, Y. He, and Y. Liu, “Research and implementation of intelligent question answering system based on knowledge Graph of traditional Chinese medicine”, Chinese Control Conference, pp. 4266-4272, 2020.

[34] D. Wang, and Y. Chen, “A novel cascade hybrid many-objective recommendation algorithm incorporating multistakeholder concerns”, Information Sciences, 2021.

[35] N. O. Dorodnykh, and A. Y. Yurin, “Knowledge Graph Augmentation Based on Tabular Data: A Case Study for Industrial Safety Inspection”, Proceedings of the Sixth International Scientific Conference, pp. 314-324, 2022.

[36] W. Fang, L. Ma, et al, “Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology”, Automation in Construction, 2020.

[37] C. Liu, and S. Yang, “Using text mining to establish knowledge graph from accident/incident reports in risk assessment”, ESWA, 2022.

[38] X. Duan, et al, “Safety-Critical Components Analysis Using Knowledge Graph For CNC Machine”, CASE, pp. 1617-1621, 2022.

[39] E. Muñoz, et al, “Facilitating prediction of adverse drug reactions by using knowledge graphs and multi-label learning models”, Briefings in bioinformatics, Vol. 20, No. 1, pp. 130-202, 2019.

[40] M. Liu, et al, “A knowledge graph-based data representation approach for IIoT-enabled cognitive manufacturing”, Advanced Engineering Informatics, Vol. 51, 2022.

[41] L. Guo, et al, “An automatic machining process decision-making system based on knowledge graph”, International Journal of Computer Integrated Manufacturing, Vol. 34, No. 12, pp. 1348-1369, 2021.

[42] H. Xu, et al, “Application of ontology and multivariate decision diagram in cloud monitor systems”, International Journal of Performability Engineering, Vol. 15, 2019.

[43] S. Namany, et al, “Sustainable energy, water and food nexus systems: A focused review of decision-making tools for efficient resource management and governance”, Journal of Cleaner Production, pp. 610-626, 2019.

[44] M. Adil, et al, “HOPCTP: a robust channel categorization data preservation scheme for industrial healthcare internet of things”, TII, Vol. 18, No. 10, pp. 7151-7161, 2022.

[45] J. Lee, B. Bagheri, and H. A. Kao, “A cyber-physical systems architecture for industry 4.0-based manufacturing systems”, Manufacturing letters, Vol. 3, pp. 18-23, 2015.

[46] B. Zhou, J. Bao, J. Li, Y. Lu, T. Liu, et al, “A novel knowledge graph-based optimization approach for resource allocation in discrete manufacturing workshops”, Robotics and Computer-Integrated Manufacturing, 2021.

[47] M. Wu, M. Hughes, S. Parbhoo, M. Zazzi, V. Roth, et al, “Beyond sparsity: Tree regularization of deep models for interpretability”, AAAI, Vol. 32, No.1, 2018.

[48] C. P. Pramod, and G. N. Pillai, “K-Means clustering based Extreme Learning ANFIS with improved interpretability for regression problems”, KBS, 2021.