A Deep-Learning-Based Use-Case Recommendation Framework for Industrial Data Analytics in Cybersecurity

RAY-I CHANG, DONG-MING HSIEH, WEL-JIE LIAO
Department of Engineering Science and Ocean Engineering
National Taiwan University
No. 1, Sec. 4, Roosevelt Rd., Da'an Dist., Taipei 106319
TAIWAN

Abstract: - Industrial cybersecurity analysis often relies on textual incident descriptions that vary widely in structure, terminology, and level of technical detail. In practical manufacturing environments, the manual interpretation of such problem narratives by domain experts becomes increasingly difficult to sustain as system complexity and data volume grow in Cyber-Physical Systems (CPS). To address this issue, this study examines the use of a deep-learning-based framework, referred to as Deep-Learning-Based Use-Case Recommendation (DLUR), for mapping unstructured industrial problem descriptions to relevant analytical use-cases. The framework combines Doc2Vec-based document representations with Bidirectional Long Short-Term Memory (BiLSTM) models to capture contextual semantic information beyond keyword-level similarity. The resulting representations are further associated with a CPS-oriented hierarchical knowledge structure, enabling recommendations to be interpreted in relation to physical assets and security contexts. Experimental evaluation was conducted using a dataset of 3,548 industrial technical documents. The results indicate that DLUR provides more stable recommendation performance than conventional Support Vector Machine and Random Forest baselines under the examined conditions. These observations suggest that the proposed approach can support analysts during the early stages of cybersecurity assessment, particularly when prior domain experience is limited.

Key-Words: - Recommendation System, Deep Learning, Long Short-Term Memory, Cyber-Physical System, Use-Case, Natural Language Processing, Cybersecurity, Digital Twins.

Received: June 23, 2025. Revised: July 25, 2025. Accepted: August 26, 2025. Published: December 31, 2025.

1 Introduction

With the increasing deployment of connected sensors, networked controllers, and data-driven monitoring systems in modern manufacturing environments, cybersecurity-related data analytics have become an integral part of daily industrial operations. Modern manufacturing systems generate large volumes of heterogeneous data while simultaneously facing increasing cybersecurity threats [1], [2], [3], [4]. Machine learning and deep learning techniques have been widely applied in analytics and security-related tasks; however, determining which analytical methods and security strategies are appropriate for a given industrial problem still relies heavily on domain expertise [5], [6].

In practice, industrial practitioners frequently describe problems using natural language rather than formal models. Bridging the gap between raw textual reports and actionable cybersecurity frameworks represents a complex task in technical translation. While use-cases provide an essential medium for mapping localized issues to reusable analytical templates, their manual identification remains inefficient. This human-centric approach is often bottlenecked by subjective interpretation and excessive processing time, necessitating a more automated, systematic alternative.

To reduce the dependence on manual expert judgment in this process, this study develops an automated use-case recommendation framework, referred to as DLUR, which leverages deep learning and natural language processing techniques. The primary contributions of this study are summarized as follows: (1) formulation of the use-case recommendation problem for industrial data analytics in cybersecurity; (2) development of a deep-learning-based recommendation framework using Doc2Vec, Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM); (3) integration of a knowledge organization mechanism with a cyber-physical system (CPS); and (4) experimental validation demonstrating the effectiveness of the proposed approach.

2 Problem Formulation

2.1 Background and Motivation

In industrial cybersecurity practice, analysts are often required to interpret a wide range of textual materials, including equipment logs, maintenance records, and incident descriptions that differ substantially in structure and terminology. As industrial systems expand in scale and interconnection, identifying analytical strategies that are both technically relevant and operationally feasible becomes increasingly difficult when such information is processed manually [7], [8], [9].

A major source of this difficulty lies in the heterogeneous and unstructured nature of problem narratives, where relevant technical cues may be expressed using inconsistent or domain-specific language. In many cases, these characteristics limit the effectiveness of rule-based or keyword-driven automation. Consequently, there is a practical need for learning-based approaches that can associate unstructured textual descriptions with established industrial use-cases in a more systematic manner.

In this study, we assume that industrial problem descriptions contain implicit semantic information related to operational conditions and data characteristics, which can be exploited by learning-based models to infer appropriate use-cases based on historical document–label relationships.

2.2 Use-Case Recommendation

To establish a rigorous basis for the recommendation task, we define a document corpus

$$D = \{d1, d2, \dots, dN\}$$

where each element d_i represents a discrete textual narrative of an industrial anomaly or operational requirement. These documents are associated with a predefined set of use-cases [10], [11]:

$$U = \{u1, u2, \dots, uM\},$$

encompassing specific analytical methodologies and cybersecurity protocols. Given a stochastic query q—provided by a user in natural language—the objective is to identify an optimal subset.

$$Uq \subseteq U$$

that maximizes the semantic alignment with the operational context of q. This process is modeled via a mapping function:

$$f: q \to Uq,$$

which is optimized through historical document-to-use-case associations.

2.3 Semantic Mapping

A primary challenge in this formulation is the inherent incommensurability of raw textual data for direct algebraic comparison. Consequently, it is necessary to project both the query q and the corpus di into a continuous, high-dimensional latent semantic space [12], [13], [14].

$$Let\ V_q \in R^K and\ V_{di} \in R^k$$

denote the k-dimensional vector representations of the query and document , respectively. In this transformed space, similarity is no longer restricted to surface-level lexical matching but is derived from deeper semantic proximity.

The framework must therefore learn a representation that encapsulates multifaceted domain nuances, including equipment functional states, telemetry characteristics, and threat vectors.

2.4 Classification-Oriented Formulation

From a machine learning perspective, we conceptualize the recommendation problem as a multi-label classification task [15]. Given an input vector *vq*, the model estimates the probability distribution:

$$P(u_j \mid q),\ u_j \in U,$$

This distribution signifies the relative relevance of each use-case to the identified problem.

The final set of recommended solutions is determined by an ordinal ranking of these probabilities, selecting the top-K candidates:

$$u_q = arg \max_{u_j \in u} P(u_j \mid q).$$

This formulation enables scalable inference and supports automated decision-making in industrial data analytics and cybersecurity applications [16].

2.5 Problem Characteristics and Challenges

The execution of this recommendation framework is impeded by several intrinsic challenges:

1. Unstructured and heterogeneous input: Problem descriptions vary in length, vocabulary, and abstraction level.
2. Domain-specific semantics: Industrial and cybersecurity terminology is highly contextual and equipment-dependent.
3. Limited labeled data: High-quality document–use-case associations are costly to obtain.
4. Interpretability requirements: Recommended use-cases must be understandable and justifiable to practitioners.

These characteristics motivate the adoption of deep learning models capable of capturing long-term dependencies and semantic structures in textual data.

3 Problem Solution

3.1 Overall Framework Architecture

To operationalize the problem formulation established in Section 2, this study proposes a deep-learning-based framework that systematically transforms unstructured industrial problem descriptions into actionable use-case recommendations. The overall architecture of the proposed DLUR framework is illustrated conceptually in Figure 1 and consists of four major functional layers: data acquisition, text representation, deep learning inference, and use-case recommendation.

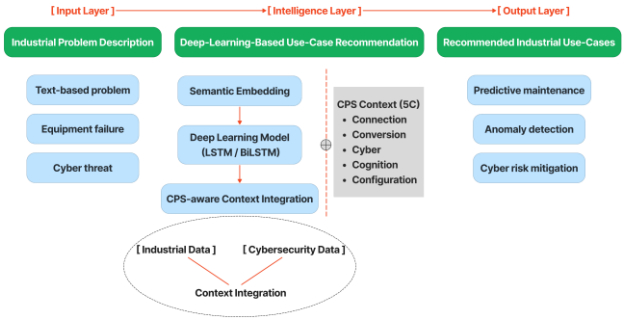

Fig. 1. Overall architecture of the proposed DLUR framework integrating industrial data analytics and cybersecurity under a CPS paradigm.
Source: Created by the authors.

At the data acquisition layer, industrial documents and user-provided problem descriptions are collected from heterogeneous sources, including technical reports, academic literature, and operational logs. These textual inputs serve as the foundation for subsequent semantic analysis.

The text representation layer is responsible for converting raw textual data into numerical vector representations. By embedding problem descriptions and historical documents into a shared latent space, semantic relationships among industrial concepts can be preserved and exploited.

The deep learning inference layer learns the mapping between text representations and predefined use-case categories. In this layer, sequential information from problem descriptions is modeled to preserve contextual relationships across sentences, which is particularly relevant when industrial incidents involve multiple interacting components or processes.

The final recommendation stage produces an ordered list of candidate use-cases based on their estimated relevance, allowing users to review and compare alternative analytical options rather than relying on a single automated decision. From an implementation perspective, the framework is organized into distinct processing stages so that changes in data representation or application scope can be accommodated without retraining the entire system.

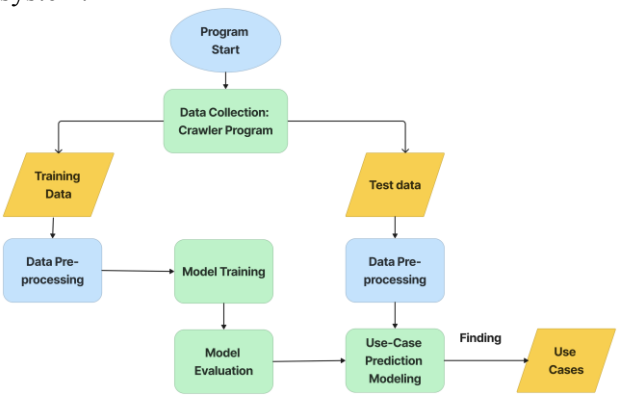

Fig. 2. Workflow of the proposed DLUR method.
Source: Created by the authors.

3.2 Overall Workflow

The overall workflow of the proposed use-case recommendation process is further illustrated in Figure 2, highlighting the sequential interactions among text preprocessing, deep learning inference, CPS-oriented knowledge integration, and use-case ranking.

3.3 Text Preprocessing and Representation

Doc2Vec is adopted in this work to obtain fixed-length representations for problem descriptions of varying length, while avoiding excessive vocabulary engineering commonly required by keyword-based methods. Before model inference, textual data are preprocessed to enhance representation quality and reduce noise. The preprocessing steps include tokenization, stop-word removal, normalization, and lemmatization. Domain-specific terms related to industrial equipment, operational states, and cybersecurity events are preserved to maintain semantic fidelity.

To represent textual inputs in a continuous semantic space, this study adopts the Doc2Vec model, which extends word embedding techniques to paragraph-level representations. Compared with traditional bag-of-words and TF–IDF approaches, Doc2Vec effectively captures contextual and semantic relationships within long and complex industrial documents [17].

Each problem description and document is mapped to a fixed-length vector, enabling consistent input to deep learning models. This representation serves as the basis for downstream learning and inference tasks.

3.4 Deep Learning Inference Models

To decode the intricate sequential dependencies within technical narratives, the framework utilizes BiLSTM networks [18], [19], [20]. In this work, bidirectional sequence modeling is used to incorporate information from both preceding and subsequent textual context, which is useful when relationships between operational events are expressed across different parts of a problem description. This allows the model to associate early indications, such as equipment behavior changes, with later references to security-related outcomes within the same narrative. The resulting latent feature representation is then funneled into a classification layer to compute the relevance probabilities of candidate use-cases.

The use of deep recurrent architectures allows the framework to generalize across diverse industrial domains and adapt to varying problem description styles.

3.5 CPS-oriented Knowledge Integration

To enhance interpretability and practical relevance, the proposed framework integrates a CPS-oriented knowledge organization mechanism [21]. In the proposed framework, industrial use-cases are organized with reference to CPS-related aspects such as equipment characteristics, data collection mechanisms, analytical processing, and security considerations, as illustrated in Figure 3.

This organization allows the outputs of the learning models to be interpreted in relation to specific system components, rather than as standalone textual matches. As a result, recommended use-cases can be examined in terms of their applicability to particular equipment configurations and operational security settings encountered in practice.

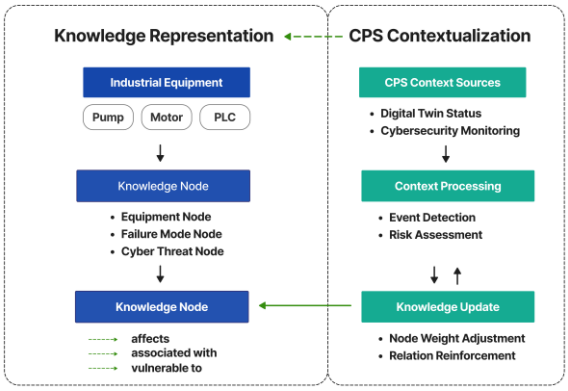

Fig. 3. CPS-oriented knowledge representation integrating industrial equipment and cybersecurity contexts for use-case recommendation.

Source: Created by the authors.

3.6 Use-Case Ranking and Recommendation

Based on the output probability distribution generated by the deep learning models, candidate use-cases are ranked according to their relevance scores. The top-K use-cases are selected as recommendations and presented to the user.

This ranking-based strategy supports flexible decision-making by allowing users to explore multiple solution alternatives. It also enables the framework to function as a decision-support system rather than a deterministic classifier, which is particularly suitable for complex industrial environments [22]. A ranking-based output is preferred over hard decision rules to allow practitioners to compare alternative solutions and incorporate domain judgment when deploying the recommended use-cases.

4 Experimental Setup

4.1 Dataset Description

To evaluate the effectiveness of the proposed DLUR framework, experiments were conducted using a dataset composed of text-based industrial and technical documents related to data analytics in cybersecurity. The documents were collected from publicly available academic repositories and technical databases, including representative industrial case studies and scholarly publications.

An overview of the dataset characteristics is summarized in Table 1.

Table 1. Overview of the datasets used for training and evaluating the proposed DLUR framework.
Source: Created by the authors.

Category	Description
Data source	Industrial problem descriptions collected from real-world industrial analytics and cybersecurity use-case documentation
Data type	Text-based descriptions of industrial operational issues, analytics requirements, and cybersecurity concerns
Application domains	Industrial data analytics, CPS, and industrial cybersecurity
Number of samples	3548 (total number of problem descriptions)
Data split strategy	k-fold cross-validation
Preprocessing steps	Text cleaning, tokenization, stop-word removal, and semantic embedding
Output labels	Recommended industrial use-cases

Each document in the dataset is associated with one or more predefined use-cases, representing analytical methods, cybersecurity strategies, or their

combinations. These use cases serve as ground-truth labels for supervised learning and performance evaluation. The dataset spans multiple industrial domains, enabling a comprehensive assessment of the proposed framework's generalization capability across diverse application scenarios.

To ensure robust and unbiased evaluation, k-fold cross-validation was adopted throughout the experimental process. The distribution of use-case categories was preserved across folds to mitigate the impact of class imbalance and to ensure consistent performance assessment across different data partitions. These problem descriptions were collected from historical industrial project reports and technical documentation, and minor inconsistencies in terminology were manually normalized during preprocessing.

4.2 Data Preprocessing

The preprocessing pipeline follows the standardized procedure detailed in Section 3.2 to maintain consistency across training and inference phases. Post-cleaning, the textual data is transformed into fixed-length, 100-dimensional vectors using the Doc2Vec model. This embedding strategy is pivotal as it allows the framework to digest variable-length incident reports while preserving the dense semantic features necessary for distinguishing between subtle cybersecurity threats.

Based on the extracted vector representations, multiple deep learning and baseline models were implemented for performance comparison. An overview of the proposed model and baseline configurations is summarized in Table 2.

Table 2. Summary of the proposed BiLSTM-based model and baseline methods used for performance comparison.
Source: Created by the authors.

Model	Description
TF-IDF + SVM	Traditional text-based classification baseline
Doc2Vec + RF	Semantic embedding-based machine learning baseline
LSTM	Deep learning baseline
BiLSTM	Proposed DLUR model

4.3 Model Training Configuration

The LSTM- and BiLSTM-based models were trained in a supervised manner using labeled associations between problem descriptions and use-cases. During training, model parameters were updated to reduce the discrepancy between predicted outputs and observed labels, as measured by categorical cross-entropy loss.

In practice, the Adam optimizer was adopted because it provided stable training behavior across different experimental runs without extensive manual tuning. Given the limited size of available industrial datasets, an early stopping strategy based on validation loss was employed to prevent unnecessary training iterations once performance improvements became marginal.

4.4 Evaluation Metrics

Model performance was examined using a set of commonly used classification metrics to facilitate comparison with existing approaches in related studies. Accuracy, precision, recall, and F1-score were reported to reflect different aspects of recommendation quality [23].

Rather than focusing on metric definitions, these measures were used to compare how consistently the framework identifies relevant use-cases and how effectively it avoids irrelevant recommendations. For experiments involving multiple use-case categories, macro-averaged scores were computed to reduce the influence of uneven label distributions across classes.

4.5 Baseline Models for Comparison

To validate the effectiveness of the proposed deep learning approach, several baseline methods were implemented for comparison. These baselines include traditional machine learning classifiers and simpler neural network architectures that rely on the same input representations [24], [25], [26].

By comparing the proposed LSTM and BiLSTM models with these baselines under identical experimental settings, the contribution of deep sequential modeling to use-case recommendation performance can be systematically evaluated.

4.6 Experimental Environment

All experiments were conducted on a standard computing platform using widely adopted deep learning frameworks, ensuring reproducibility and extensibility.

5 Experimental Results

5.1 Overall Performance Comparison

This section presents the experimental results of the proposed DLUR framework. The performance of the LSTM- and BiLSTM-based models was examined under the same experimental settings as the baseline approaches described in Section 4. A summary of the quantitative evaluation results is provided in Table 3.

The comparison shows that learning-based models generally achieve higher accuracy and F1-scores than

the selected baseline methods, although the extent of improvement differs across use-case categories [27], [28], [29]. Among the evaluated architectures, the BiLSTM model yields the highest average performance on several metrics, particularly in scenarios where contextual information is distributed across different parts of the problem description.

Table 3. Quantitative performance comparison of the proposed BiLSTM-based model and baseline methods using k-fold cross-validation.
Source: Created by the authors.

Method	Precision	Recall	F1-score
TF-IDF + SVM	0.71	0.68	0.69
Doc2Vec + Random Forest	0.76	0.74	0.75
LSTM	0.83	0.81	0.82
BiLSTM	0.88	0.86	0.87

As illustrated in Figure 4, performance variations can be observed among different model configurations, indicating that certain types of industrial problems remain more difficult to characterize using textual information alone.

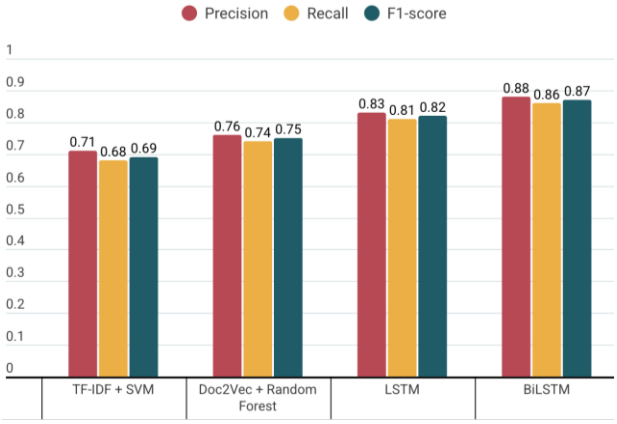

Fig. 4. Performance comparison of deep learning-based and traditional baseline methods for industrial use-case recommendation in terms of precision, recall, and F1-score.
Source: Created by the authors.

5.2 Effect of Deep Sequential Modeling

A granular analysis reveals that the BiLSTM's success is rooted in its ability to leverage bidirectional contextual dependencies. Industrial reports often contain distal dependencies where the initial description of a physical anomaly (e.g., a pressure drop) is only contextualized by a later mention of a specific network protocol.

While the standard LSTM captures forward-moving sequences, the BiLSTM's dual-processing layers allow it to synthesize these cues more holistically, resulting in a more discriminative feature representation.

5.3 Impact of Text Representation Strategy

The adoption of Doc2Vec proved to be a critical design choice. Traditional keyword-based methods like TF-IDF often fail to account for the synonymy and polysemy prevalent in engineering jargon.

In contrast, Doc2Vec embeds entire problem descriptions into a continuous latent space, preserving the semantic "spirit" of the report rather than just its surface-level vocabulary. This provided a more stable and noise-resistant input for our deep learning inference layer.

5.4 Comparison with Baseline Methods

The SVM and Random Forest baselines show lower Recall values in this evaluation, particularly when problem descriptions are abstract or expressed using heterogeneous terminology. This pattern suggests that these models may have difficulty capturing relevant information in such cases.

In comparison, the learning-based framework exhibits more consistent performance across different industrial scenarios, especially when problem descriptions involve multiple interacting operational factors.

5.5 Discussion on Practical Applicability

From a deployment perspective, the DLUR framework significantly mitigates the cognitive load placed on cybersecurity analysts. By providing a ranked list of relevant use-cases, the system functions as a robust decision-support engine, facilitating rapid intervention even for novice practitioners.

Furthermore, the integration of the CPS-centric hierarchy ensures that the model's "black-box" predictions are aligned with physical equipment states, enhancing both the interpretability and the operational feasibility of the recommendations.

5.6 Limitations

Despite its promising performance, the proposed framework has several limitations. First, the quality of recommendations is dependent on the availability and diversity of labeled industrial documents. Limited coverage of certain domains may affect generalization performance.

Second, while deep learning models provide strong predictive capabilities, their internal decision-making processes remain relatively opaque. Although a CPS-oriented organization partially mitigates this issue, further work on explainable AI techniques is necessary to enhance user trust.

6 Conclusion and Future work

6.1 Conclusion

This study establishes the DLUR framework as a viable computational bridge between unstructured industrial narratives and specialized cybersecurity solutions. By formalizing the recommendation task as a semantic mapping problem, we have successfully addressed the persistent gap between initial problem identification and the selection of operationally viable analytical patterns. The synergy between Doc2Vec embeddings and BiLSTM networks proved instrumental in decoding the bidirectional contextual dependencies that are often obscured in complex technical reports.

Furthermore, the integration of a CPS-oriented knowledge hierarchy ensures that the model's outputs are not merely statistical artifacts, but rather actionable guidance aligned with physical asset configurations.

6.2 Implications for Industrial Applications

From an operational standpoint, our framework introduces a systematic methodology for identifying cybersecurity strategies, effectively lowering the cognitive threshold for novice engineers and cross-domain teams. By automating the interpretation of natural language descriptions, the proposed system facilitates rapid knowledge transfer within organizations undergoing digital transformation, thereby reducing the dependency on scarce human expertise.

6.3 Limitations and Future Work

Attaining high predictive accuracy does not imply that the framework is without its constraints. A primary bottleneck remains the dependency on the fidelity of the labeled corpus; any significant underrepresentation of niche industrial domains could potentially skew the model's generalization capabilities. Additionally, the intrinsic "black-box" nature of deep learning inference layers presents an interpretability hurdle that must be resolved to secure full trust in safety-critical industrial settings.

To circumvent these hurdles, our future research will pivot toward:

1. Explainable AI (XAI) Integration: Developing transparency layers to elucidate the causal logic behind each recommendation.
2. Hybrid Knowledge Modeling: Melding deep learning with knowledge graphs or causal reasoning to enhance precision in non-linear CPS scenarios.
3. Cross-Domain Data Augmentation: Leveraging larger, more heterogeneous datasets to strengthen the model's robustness against domain shift.

6.5 Final Remarks

In summary, this research underscores that deep-learning-driven semantic modeling is no longer a theoretical luxury but a practical necessity for modern industrial cybersecurity . By establishing a systematic pipeline from raw incident reports to structured use-case knowledge, the DLUR framework provides a robust and extensible foundation for supporting data-driven decision-making in the increasingly complex landscape of modern CPS.

References:

[1] H. Kagermann, W. Wahlster, and J. Helbig, Recommendations for Implementing the Strategic Initiative Industry 4.0, Acatech, Germany, 2013.

[2] E. A. Lee, "Cyber-physical systems: Design challenges," Proceedings of the IEEE International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing (ISORC), Orlando, FL, USA, pp. 363–369, 2008, doi: 10.1109/ISORC.2008.25.

[3] 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, doi: 10.1016/j.mfglet.2014.12.001.

[4] D. Abraham, G. Erceylan, V. Gkioulos, and S. Katsikas, "Towards enhanced cybersecurity in industrial control systems: A systematic review of context-based modeling, digital twins, and machine learning approaches," International Journal of Information Security, vol. 24, p. 243, 2025, doi: 10.1007/s10207-025-01158-1.

[5] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital twin: Enabling technologies, challenges, and open research," IEEE Access, vol. 8, pp. 108952–108971, 2020, doi: 10.1109/ACCESS.2020.2998358.

[6] F. Tao, Q. Qi, L. Wang, and A. Nee, "Digital twin-driven smart manufacturing," IEEE Transactions on Industrial Informatics, vol. 15, no. 4, pp. 2405–2416, 2019, doi: 10.1109/TII.2018.2873186.

[7] S. Mittal and O. P. Sangwan, "Big data analytics using data mining techniques: A survey," in Advanced Informatics for Computing Research,

Communications in Computer and Information Science, vol. 955, Springer, Singapore, pp. 243–254, 2019, doi: 10.1007/978-981-13-3140-4_24.

[8] N. Ur-Rahman and J. A. Harding, "Textual data mining for industrial knowledge management and text classification: A business-oriented approach," Expert Systems with Applications, vol. 39, no. 5, pp. 4729–4739, 2012, doi: 10.1016/j.eswa.2011.09.124.

[9] M. M. Aslam, A. Tufail, and M. N. Irshad, "Survey of deep learning approaches for securing industrial control systems: A comparative analysis," Cyber Security and Applications, vol. 3, Art. no. 100096, 2025, doi: 10.1016/j.csa.2025.100096.

[10] F. Ricci, L. Rokach, and B. Shapira, Recommender Systems Handbook, Springer, New York, USA, 2015.

[11] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," Proceedings of the 26th International World Wide Web Conference (WWW), Perth, Australia, pp. 173–182, 2017, doi: 10.1145/3038912.3052569.

[12] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," Proceedings of the International Conference on Learning Representations (ICLR), 2013, doi: 10.48550/arXiv.1301.3781.

[13] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global vectors for word representation," Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, pp. 1532–1543, 2014, doi: 10.3115/v1/D14-1162.

[14] A. Berenguer, J.-N. Mazón, and D. Tomás, "Word embeddings for retrieving tabular data from research publications," Machine Learning, vol. 113, pp. 2227–2248, 2024, doi: 10.1007/s10994-023-06472-0.

[15] D. Zhang, S. Zhao, Z. Duan, J. Chen, Y. Zhang, and J. Tang, "A multi-label classification method using a hierarchical and transparent representation for paper-reviewer recommendation," ACM Transactions on Information Systems, vol. 38, no. 1, Art. no. 5, 2020, doi: 10.1145/3361719.

[16] T.-Y. Liu, Learning to Rank for Information Retrieval, Springer, Berlin, Germany, 2011.

[17] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," Proceedings of the 31st International Conference on Machine Learning (ICML), Beijing, China, pp. 1188–1196, 2014.

[18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[19] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," IEEE Tran sactions on Signal Processing, vol. 45, no. 11, pp. 2673–2681, 1997, doi: 10.1109/78.650093.

[20] Y. Liu, Y. Dai, and T. Bakhshi, "Deep learning in cybersecurity: A hybrid BERT–LSTM network for SQL injection attack detection," IET Information Security, 2024, doi: 10.1049/2024/5565950.

[21] J. Lee, E. Lapira, B. Bagheri, and H.-A. Kao, "Recent advances and trends in predictive manufacturing systems in big data environments," Manufacturing Letters, vol. 1, pp. 38–41, 2013, doi: 10.1016/j.mfglet.2013.09.005.

[22] C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval, Cambridge University Press, Cambridge, UK, 2008.

[23] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," Journal of Machine Learning Technologies, vol. 2, no. 1, pp. 37–63, 2011, doi: 10.9735/2229-3981.

[24] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features," Proceedings of the European Conference on Machine Learning (ECML), Chemnitz, Germany, pp. 137–142, 1998.

[25] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, pp. 273–297, 1995.

[26] A. Alharthi, M. Alaryani, and S. Kaddoura, "A comparative study of machine learning and deep learning models in binary and multiclass classification for intrusion detection systems," Array, vol. 26, Art. no. 100406, 2025, doi: 10.1016/j.array.2025.100406.

[27] Y. Kim, "Convolutional neural networks for sentence classification," Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, pp. 1746–1751, 2014, doi: 10.3115/v1/D14-1181.

[28] R. Johnson and T. Zhang, "Deep pyramid convolutional neural networks for text categorization," Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL), Vancouver, Canada, pp. 562–570, 2017, doi: 10.18653/v1/P17-1052.

[29] E. Iturbe, C. Dalamagkas, P. Radoglou-Grammatikis, E. Rios, and N. Toledo, "A pattern-aware LSTM-based approach for APT detection leveraging a realistic dataset for critical infrastructure security," Future Generation Computer Systems, vol. 178, Art. no. 108308, 2026, doi: 10.1016/j.future.2025.108308.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors contributed equally to this work at all stages, including problem formulation, methodology design, experimental analysis, and manuscript preparation.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare.