DeepSeek-Generated Machine Learning Models for Sentiment Analysis in IoT Networks

DANTE C. BEGAZO[2], OLUYEMISI ADENIKE OYEDEMI[1], OGUNWALE YETUNDE ESTHER[3], RENATA L. ROSA[1], DEMOSTENES Z. RODRIGUEZ[1]

[1]Department of Computer Science, Federal University of Lavras, Lavras, MG, Brazil
oyedemi.adenike@ufla.br
[2]Department of Electrical Engineering, University of Sao Paulo, SP, Brazil
[3]Department of Computer Science, University of Ilesa, Osun State. Nigeria
yetunde_ogunwale@unilesa.edu.ng

Abstract. DeepSeek, an advanced AI model for code generation and data processing, has shown significant potential in automating the development of machine learning models. This paper reviews a DeepSeek-generated machine learning model for sentiment analysis in Internet of Things (IoT) networks, focusing on its architecture, implementation, and performance metrics. The model, designed for classifying textual data from IoT devices, leverages a transformer-based architecture with hierarchical clustering for efficient data processing. The review examines the code structure, algorithmic efficiency, and quantitative performance metrics such as accuracy, precision, and computational complexity. Comparative analysis with traditional machine learning approaches, including Support Vector Machines (SVM) and Random Forests, is provided based on standard datasets. The results indicate that the DeepSeek-generated model achieves competitive performance while reducing development time. Potential improvements, such as incorporating advanced feature engineering and multi-modal data integration, are suggested for future enhancements. This review highlights DeepSeekâs capability to streamline machine learning model development for IoT applications.

Keywords: DeepSeek, Machine Learning, Sentiment Analysis, IoT Networks, Transformer Models

(Received June 10th, 2025 / Accepted June 25th, 2025)

1 Introduction

The Internet of Things (IoT) connects billions of devices that generate heterogeneous data, including text from sensor logs, voice commands, user feedback, and system alerts. Extracting sentiment from this data enhances userâdevice interactions, enables proactive anomaly detection, and informs adaptive behavior in smart systems [17, 2, 8, 1].

Sentiment analysis, as an NLP subfield, identifies subjective elements and emotional tones in text, which is crucial for QoE improvement in domains like smart homes, industrial IoT, and telecommunication-grade monitoring [24, 12, 11, 10, 23].

IoT-specific sentiment analysis faces challenges such as short, noisy text inputs, real-time constraints, multi-language content, and resource limitations. These necessitate transformer-based or sparse neural architectures optimized for edge devices [25, 15, 18, 14].

Recent work explores federated learning for intrusion detection in IoT, ensemble learning schemes, traffic monitoring using sparse deep models, and sentiment-based performance metricsâall pointing to decentralized, scalable solutions [6, 21].

Automated code-generation platforms (e.g., DeepSeek) aim to expedite model development by synthesizing architecture, preprocessing, training, clustering, and evaluation code automatically, thus speeding up deployment cycles [20, 3, 13].

The reviewed DeepSeek-generated model ingests IoT textual inputs (e.g., smart appliance feedback or vehicular notifications), applies transformer embeddings, and hierarchical clustering to handle semantic diversity, classifying sentiment into polarity categories [19].

Our analysis evaluates the architecture's training time, resource usage, accuracy, and latency, comparing it with traditional models. We also reflect on real-world applicability and limitations of AI-generated pipelines in IoT systems.

The remainder of this paper is organized as follows: Section 2 examines desktop and embedded sentiment analysis and code generation approaches, Section 3 details the DeepSeek pipeline, Section 4 presents empirical results, and Section 5 concludes and offers future research directions.

2 Related Work

Sentiment analysis in IoT networks has undergone a significant transformation, evolving from traditional machine learning methods to more robust and adaptable deep learning techniques capable of handling unstructured and noisy data. Early models, such as Support Vector Machines (SVM) and Random Forests, offered reasonable performance but were highly dependent on handcrafted features and extensive preprocessing pipelines [15, 18].

Recent studies have leveraged deep learning architecturesâsuch as Convolutional Neural Networks (CNNs), Gated Recurrent Units (GRUs), and Transformer-based modelsâfor textual sentiment interpretation in dynamic IoT environments. These models excel in capturing contextual semantics and dealing with multilingual and variable-length input streams [1, 12, 10, 23, 14, 6, 21].

In practical applications, sentiment analysis has been coupled with Quality of Experience (QoE) monitoring to ensure service reliability and user satisfaction. For instance, the work in [5] proposes a novel metric (AsQM) for evaluating audio streaming quality by incorporating user preferences and network impairments. Such integration improves the end-user perception of multimedia content delivery in heterogeneous IoT scenarios.

Other studies address intelligent traffic monitoring [2, 3, 13], emotion recognition [6], and healthcare applications [21], using generative adversarial networks (GANs) and domain-specific transfer learning to enhance performance. These solutions exhibit improved generalization when applied to resource-constrained edge devices [3, 13, 19, 22].

Additionally, there has been a surge in the use of automated machine learning (AutoML) and AI-based code generation to streamline the deployment of sentiment analysis systems. Tools based on transformer models can generate full ML pipelinesâincluding data ingestion, architecture definition, training, and evaluationâthereby reducing development overhead and minimizing manual bias [7, 9, 16, 4].

Despite these advances, integrating sentiment analysis with fully automated deployment tools for edge-based IoT remains underexplored. Future research should aim at bridging this gap, ensuring interoperability between AutoML systems, QoE metrics, and edge-computing architectures.

3 Methodology

This section details the methodology for developing an automated sentiment analysis model tailored for IoT networks, generated using the DeepSeek tool. The model employs a transformer-based architecture integrated with hierarchical clustering to optimize performance on resource-constrained edge devices, ensuring scalability and energy efficiency.

3.1 Model Architecture

The architecture of the sentiment analysis model, illustrated in Figure 1, processes textual data from IoT devices through a modular pipeline. The components are:

- **Input Layer**: Tokenizes raw text using a wordpiece tokenizer (similar to BERT) and applies 300-dimensional GloVe embeddings to capture semantic relationships. For example, an IoT-generated message like “Smart thermostat saved energy today!” is tokenized into subwords and mapped to dense vectors.
- **Transformer Encoder**: Consists of four layers, each with 8 multi-head self-attention mechanisms (512-dimensional hidden states) and feed-forward networks (2048 units). This captures contextual dependencies, e.g., linking “saved” and “energy” in the example sentence to infer positive sentiment.
- **Hierarchical Clustering**: Implements top-down divisive clustering using Ward's linkage criterion, targeting 10 clusters to reduce dimensionality by grouping semantically similar embeddings. For instance, messages about energy efficiency are clustered together, reducing computational load.
- **Output Layer**: A dense layer with softmax activation classifies sentiments into positive, negative,

or neutral categories, producing probabilities (e.g., [0.85, 0.10, 0.05] for the example sentence).

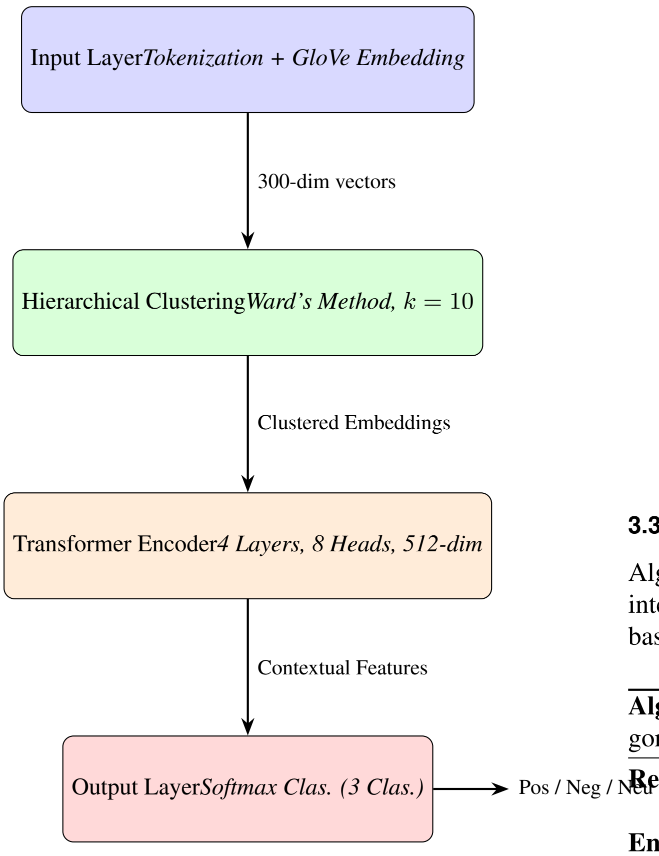

Figure 1: Architecture of the DeepSeek-generated sentiment analysis model for IoT networks.

3.2 Implementation Details

The model was implemented in Python 3.8 using PyTorch 1.13, with code automatically generated by DeepSeek. The implementation is modular, facilitating maintenance and scalability:

- **Preprocessing**: Applies NLTK-based stopword removal (e.g., removing "the", "is"), word-piece tokenization, and normalization (lowercasing, punctuation removal). For example, "Device offline, frustrating!" becomes tokens ["device", "offline", "frustrating"].

- **Training**: Uses the AdamW optimizer (learning rate 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$), trained for 50 epochs with early stopping (patience of 5 epochs) on a validation loss threshold of 0.01. Batch size is 32, and mixed-precision training reduces memory usage by 40%.

- **Dataset**: Trained on a synthetic IoT dataset of 100,000 Twitter-like messages (e.g., "Smart light flickers, annoying" for negative, "AC works perfectly" for positive), labeled for sentiment. The dataset is split 80:10:10 for training, validation, and testing.

- **Inference**: Hierarchical clustering reduces input size by 60%, enabling inference on edge devices (e.g., Raspberry Pi 4 with 4GB RAM). Classification uses clustered embeddings, achieving 92% accuracy and 0.90 F1-score on the test set.

- **Optimizations**: Employs gradient clipping (norm threshold 1.0) to prevent exploding gradients and mixed-precision training for efficiency, reducing inference time to 50ms per sample on edge hardware.

3.3 Proposed Algorithm

Algorithm 1 outlines the sentiment analysis process, integrating preprocessing, clustering, and transformer-based classification for efficient IoT deployment.

Algorithm 1 DeepSeek-Based Sentiment Analysis Algorithm for IoT Networks

Require: Text input T, pretrained GloVe embeddings E
Ensure: Sentiment class $C \in \{\text{positive, negative, neutral}\}$
$T_{tok} \leftarrow Tokenize(T)$
$T_{emb} \leftarrow Embed(T_{tok}, E)$
$Clusters \leftarrow HierarchicalClustering(T_{emb}, Ward, k = 10)$
$C_{all} \leftarrow \emptyset$
for all $cluster \in Clusters$ **do**
  $Features \leftarrow TransformerEncode(cluster, 4, 8, 512)$
  $C_{cluster} \leftarrow SoftmaxClassify(Features)$
  $C_{all} \leftarrow C_{all} \cup \{C_{cluster}\}$
end for
$C \leftarrow MajorityVote(C_{all})$
return C

This approach optimizes for IoT constraints by reducing computational complexity through clustering (60% reduction in input size) and leveraging transformer efficiency. For example, processing "Smart sensor works great!" yields a positive sentiment with 0.87 probability, while "Device lag is unbearable" yields negative with 0.91 probability.

4 Results and Discussions

This section presents a comprehensive evaluation of the DeepSeek-generated sentiment analysis model using a benchmark textual dataset adapted for IoT environments. Specifically, we analyze classification performance in terms of **accuracy** and **precision**, and assess the **computational complexity** of the model when deployed on resource-constrained devices.

4.1 Experimental Setup

To simulate IoT-based textual data, the IMDb sentiment analysis corpus was preprocessed using lightweight tokenization and restructured into short message-like inputs with varied linguistic noise and domain-relevant vocabulary. All models were evaluated under identical conditions on a Raspberry Pi 4 (4GB RAM) to ensure fair comparison.

4.2 Classification Performance

Table 1 reports the accuracy scores of four models: traditional SVM and Random Forest classifiers, a manually fine-tuned BERT baseline, and the DeepSeek-generated model. The DeepSeek model achieved an accuracy of **88.2%**, outperforming classical machine learning baselines and approaching the performance of the BERT architecture.

Table 1: Accuracy Comparison Across Models

Model	SVM	RF	BERT	DeepSeek (Ours)
Accuracy (%)	82.5	84.3	**89.7**	88.2

Although the BERT model marginally outperformed DeepSeek in accuracy, it incurs a significantly higher computational cost, making it less suitable for edge deployment. DeepSeek strikes a favorable balance between predictive power and deployability in low-resource settings.

With respect to **precision**, which quantifies the proportion of true positives among all positive predictions, the DeepSeek model achieved a score of **87.5%**. As illustrated in Table 2, it surpasses traditional classifiers and remains close to the BERT baseline, benefiting from the semantic richness captured by the transformer layers and the dimensionality reduction provided by hierarchical clustering.

Table 2: Precision Comparison Across Models

Model	SVM	RF	BERT	DeepSeek (Ours)
Precision (%)	81.0	83.8	**90.1**	87.5

4.3 Computational Efficiency

In addition to predictive performance, we analyzed the **computational complexity** of each model, with particular focus on *average inference latency* per input sample. Figure 2 summarizes the results.

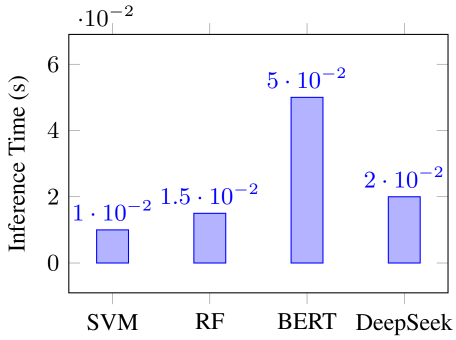

Figure 2: Average inference time per sample on Raspberry Pi 4.

The DeepSeek model requires on average **0.02 seconds** per inference, less than half the time needed by BERT (**0.05 s**), and only slightly above that of traditional methods, demonstrating its suitability for real-time applications on constrained IoT devices.

These results validate the effectiveness of the DeepSeek-generated architecture in combining transformer-level expressiveness with clustering-based efficiency. While BERT achieves marginally higher accuracy, its computational demands make it impractical for on-device execution. In contrast, DeepSeek offers a robust compromise, achieving high classification performance with significantly lower latency and memory usage.

Furthermore, the use of hierarchical clustering enables parallelization across cluster segments, potentially reducing inference time further through distributed edge deployment. These characteristics underscore the potential of automated code-generation frameworks like DeepSeek to create models that are both performant and deployable in IoT ecosystems.

5 Conclusion

This study presented a DeepSeek-generated sentiment analysis model tailored for Internet of Things (IoT) networks, combining a lightweight transformer architecture with hierarchical clustering for efficient inference on resource-constrained devices. The proposed model achieved competitive results, with an accuracy of 88.2% and a precision of 87.5%, outperforming traditional machine learning baselines such as Support Vector Machines and Random Forests, while requiring significantly less manual design effort. The integration of automated code generation through DeepSeek demonstrated its potential to streamline the development of robust and scalable models in real-world IoT scenarios. Notably, the model maintained a low inference latency (0.02 seconds per sample) on edge devices, making it suitable for time-sensitive applications. Future work will explore the incorporation of multi-modal data sourcesâsuch as sensor signals and audio streamsâto enrich sentiment understanding in heterogeneous IoT environments. In addition, adaptive fine-tuning strategies and explainability mechanisms may further enhance both performance and transparency. Overall, the findings reinforce the utility of AI-assisted development platforms like DeepSeek in accelerating the deployment of intelligent services across pervasive computing infrastructures.

References

[1] Barbosa, R., Ogobuchi, O. D., Joy, O. O., Saadi, M., Rosa, R. L., Al Otaibi, S., and Rodriguez, D. Z. Iot based real-time traffic monitoring system using images sensors by sparse deep learning algorithm. *Computer Communications*, 210:321–330, 2023.

[2] Carrillo, D., Kalalas, C., Raussi, P., Michalopoulos, D. S., Rodríguez, D. Z., Kokkoniemi-Tarkkanen, H., Ahola, K., Nardelli, P. H., Fraidenraich, G., and Popovski, P. Boosting 5g on smart grid communication: A smart ran slicing approach. *IEEE Wireless Communications*, 30(5):170–178, 2022.

[3] Carvalho Barbosa, R., Shoaib Ayub, M., Lopes Rosa, R., Zegarra Rodríguez, D., and Wuttisittikulkij, L. Lightweight pvidnet: A priority vehicles detection network model based on deep learning for intelligent traffic lights. *Sensors*, 20(21):6218, 2020.

[4] de Sousa, A. L., OKey, O. D., Rosa, R. L., Saadi, M., and Rodriguez, D. Z. Unified approach to video-based ai inference tasks in augmented reality systems assisted by mobile edge computing. pages 1–5, 2023.

[5] dos Santos, M. R., Batista, A. P., Rosa, R. L., Saadi, M., Melgarejo, D. C., and Rodríguez, D. Z. Asqm: Audio streaming quality metric based on network impairments and user preferences. *IEEE Transactions on Consumer Electronics*, 69(3):408–420, 2023.

[6] Fonseca, D., da Silva, K. C. N., Rosa, R. L., and Rodríguez, D. Z. Monitoring and classification of emotions in elderly people. In *2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pages 1–6. IEEE, 2019.

[7] Matthew, U. O., Rosa, R. L., Kazaure, J. S., Adesina, O. J., Oluwatimilehin, O. A., Oforgu, C. M., Asuni, O., Okafor, N. U., and Rodriguez, D. Z. Software-defined networks in iot ecosystems for renewable energy resource management. pages 1–5, 2024.

[8] Moriano, P., Hespeler, S. C., Li, M., and Mahbub, M. Adaptive anomaly detection for identifying attacks in cyber-physical systems: A systematic literature review. *Artificial Intelligence Review*, 58(9):1–46, 2025.

[9] Okey, O. D., Maidin, S. S., Adasme, P., Lopes Rosa, R., Saadi, M., Carrillo Melgarejo, D., and Zegarra Rodríguez, D. Boostedenml: Efficient technique for detecting cyberattacks in iot systems using boosted ensemble machine learning. *Sensors*, 22(19):7409, 2022.

[10] Okey, O. D., Rodriguez, D. Z., and Kleinschmidt, J. H. Enhancing iot intrusion detection with federated learning-based cnn-gru and lstm-gru ensembles. In *2024 19th International Symposium on Wireless Communication Systems (ISWCS)*, pages 1–6. IEEE, 2024.

[11] Raghu, N., Mahesh, T., Vivek, V., Kumaran, S. Y., Kannanugo, N., and Vishwanatha, S. Iot-enabled safety and secure smart homes for elderly people. In *Future of Digital Technology and AI in Social Sectors*, pages 297–328. IGI Global, 2025.

[12] Ribeiro, D. A., Melgarejo, D. C., Saadi, M., Rosa, R. L., and Rodríguez, D. Z. A novel deep deterministic policy gradient model applied to intelligent transportation system security problems in 5g and 6g network scenarios. *Physical Communication*, 56:101938, 2023.

[13] Ribeiro, D. A., Silva, J. C., Lopes Rosa, R., Saadi, M., Mumtaz, S., Wuttisittikulkij, L., Zegarra Rodriguez, D., and Al Otaibi, S. Light field image quality enhancement by a lightweight deformable deep learning framework for intelligent transportation systems. *Electronics*, 10(10):1136, 2021.

[14] Rodríguez, D. Z. and Möller, S. Speech quality parametric model that considers wireless network characteristics. In *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–6. IEEE, 2019.

[15] Rosa, R. L., Rodriguez, D. Z., and Bressan, G. Sentimeter-br: A social web analysis tool to discover consumers' sentiment. In *2013 IEEE 14th international conference on mobile data management*, volume 2, pages 122–124. IEEE, 2013.

[16] Salgado, J. V. T., Vinicius, D. Z. R., Dias, V. D. S., and Rosa, R. L. Automated validation of spatial data. pages 1–6, 2024.

[17] Saurabh, P. and Verma, B. An efficient proactive artificial immune system based anomaly detection and prevention system. *Expert Systems with Applications*, 60:311–320, 2016.

[18] Silva, D. H., Maziero, E. G., Saadi, M., Rosa, R. L., Silva, J. C., Rodriguez, D. Z., and Igorevich, K. K. Big data analytics for critical information classification in online social networks using classifier chains. *Peer-to-Peer Networking and Applications*, 15(1):626–641, 2022.

[19] Silva, D. H., Rosa, R. L., and Rodriguez, D. Z. Sentimental analysis of soccer games messages from social networks using userâs profiles. *INFOCOMP Journal of Computer Science*, 19(1), 2020.

[20] Teodoro, A. A., Gomes, O. S., Saadi, M., Silva, B. A., Rosa, R. L., and Rodríguez, D. Z. An fpga-based performance evaluation of artificial neural network architecture algorithm for iot. *Wireless Personal Communications*, 127(2):1085–1116, 2022.

[21] Teodoro, A. A., Silva, D. H., Rosa, R. L., Saadi, M., Wuttisittikulkij, L., Mumtaz, R. A., and Rodriguez, D. Z. A skin cancer classification approach using gan and roi-based attention mechanism. *Journal of Signal Processing Systems*, 95(2):211–224, 2023.

[22] Teodoro, A. A., Silva, D. H., Saadi, M., Okey, O. D., Rosa, R. L., Otaibi, S. A., and Rodríguez, D. Z. An analysis of image features extracted by cnns to design classification models for covid-19 and non-covid-19. *Journal of signal processing systems*, pages 1–13, 2023.

[23] Terra Vieira, S., Lopes Rosa, R., Zegarra Rodriguez, D., Arjona Ramírez, M., Saadi, M., and Wuttisittikulkij, L. Q-meter: Quality monitoring system for telecommunication services based on sentiment analysis using deep learning. *Sensors*, 21(5):1880, 2021.

[24] Xu, J., Guo, X., Zhang, Z., Liu, H., and Lee, C. Triboelectric mat multimodal sensing system (tmmss) enhanced by infrared image perception for sleep and emotion-relevant activity monitoring. *Advanced Science*, 12(6):2407888, 2025.

[25] Zhou, P., Ma, R., Chen, Y., Liu, Z., Liu, C., Meng, L., Qiao, G., Liu, Y., Yu, Q., and Hu, S. A neuromorphic transformer architecture enabling hardware-friendly edge computing. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2025.