# Unsupervised Topic Discovery and Text Extraction with Generative AI and Embedding Techniques

[1]Mr. Selvaprasanth P
[1]Assistant Professor, Department of Electronics and Communication Engineering,
Sethu Institute of Technology, Virudhunagar
[1]selvaprasanth9619@gmail.com

**Abstract:** The exponential growth of unstructured textual data across diverse domains has created an urgent need for efficient methods to extract meaningful information without relying on extensive labeled datasets. Unsupervised topic discovery and text extraction have emerged as critical tasks for summarizing, organizing, and understanding large-scale text corpora. Recent advances in generative artificial intelligence (AI) and embedding-based representation learning have transformed these tasks, enabling more accurate, scalable, and interpretable outcomes. This paper presents a comprehensive overview of the integration of generative AI models with embedding techniques to achieve unsupervised topic discovery and text extraction, highlighting the methodological innovations and practical implications. Traditional topic modeling approaches, such as Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF), rely on probabilistic frameworks that often require assumptions about the number of topics or specific data distributions. These methods may struggle with the semantic richness and contextual complexity of modern textual datasets. In contrast, embedding techniques, which represent words, sentences, or documents as dense vector representations in continuous latent spaces, capture nuanced semantic and syntactic relationships. Techniques like Word2Vec, GloVe, and more recently contextualized embeddings from transformer-based models (e.g., BERT, GPT) have demonstrated superior capability in encoding textual context, thereby enhancing downstream unsupervised learning tasks. Generative AI models, especially those based on transformer architectures, have further expanded the potential for unsupervised text analysis. Their ability to generate coherent, context-aware text representations and perform conditional generation enables innovative approaches for topic discovery. For example, generative models can be fine-tuned to produce topic-relevant summaries or simulate latent topic distributions without explicit supervision. When combined with embeddings, these models facilitate a two-step process: first, representing the textual corpus in a high-dimensional semantic space; second, clustering or extracting salient themes and key phrases based on distance metrics or similarity scores. This integration leads to several practical advances. Embedding-based clustering methods (e.g., K-means, hierarchical clustering, or density-based algorithms) applied to generative AI-derived representations allow flexible discovery of topics that better reflect semantic coherence and human interpretability. Additionally, generative models can assist in extracting representative sentences or keywords that encapsulate the essence of discovered topics, enhancing interpretability and downstream usability. Such systems are particularly valuable in domains like social media analysis, customer feedback mining, scientific literature review, and legal document summarization, where labeled data is scarce or unavailable. Furthermore, recent developments in self-supervised learning and contrastive learning paradigms have improved embedding quality, making unsupervised topic discovery more robust to noise and domain shifts. Combining these with generative pretraining yields models capable of capturing both local and global semantic patterns, which are critical for coherent topic formation and extraction. In conclusion, the synergy between generative AI and embedding techniques marks a significant advancement in unsupervised topic discovery and text extraction. This paradigm enables more adaptable, scalable, and semantically rich analysis of vast unstructured text data. Future research directions include improving interpretability through explainable AI techniques, integrating multimodal data sources, and developing domain-adaptive models to further enhance the applicability and accuracy of unsupervised text mining frameworks.

**Keywords:** Unsupervised learning, topic discovery, text extraction, generative AI, embedding techniques, transformer models, natural language processing, semantic embeddings, clustering, latent topic modeling, self-supervised learning, contrastive learning, text summarization, document clustering, vector representations, BERT, GPT, unsupervised text mining, interpretability, deep learning, natural language understanding

**Corresponding Author:** *Mr. Selvaprasanth P*
*Assistant Professor / ECE, Sethu Institute of Technology,*
*Virudhunagar , Tamil Nadu, India*
*Mail: selvaprasanth9619@gmail.com*

## INTRODUCTION

In today's data-driven world, the volume of unstructured textual information generated daily—from social media posts, customer reviews, news articles, scientific publications, to legal documents—has grown exponentially. Extracting meaningful insights from this vast amount of text is essential for decision-making, knowledge discovery, and efficient information retrieval. However, the lack of labeled data and the high cost of manual annotation pose significant challenges to traditional supervised learning approaches. This situation has spurred the development of unsupervised methods for topic discovery and text extraction, which aim to identify hidden thematic structures and extract salient content without requiring prior annotations.

Topic discovery refers to the process of uncovering latent themes or subjects within a large corpus, providing a high-level understanding of the text. Text extraction involves selecting key sentences, phrases, or keywords that succinctly summarize or represent these topics. Together, these tasks facilitate summarization, indexing, and exploration of textual data, enabling users to navigate complex datasets efficiently. Historically, unsupervised topic modeling techniques such as Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF) have been widely used. These probabilistic models rely on word co-occurrence patterns and assume specific generative processes. While effective in some contexts, these approaches often struggle to capture deep semantic relationships and context-dependent meanings, limiting their applicability for complex or noisy datasets.

The advent of deep learning and particularly the rise of transformer-based models have revolutionized natural language processing (NLP). Embedding techniques, which encode words, sentences, or entire documents as dense vectors in continuous semantic spaces, provide a powerful way to capture rich linguistic patterns and contextual nuances. Early word embedding methods like Word2Vec and GloVe represented static meanings of words but were limited in handling polysemy and context variation. More recent contextual embeddings generated by models such as BERT, RoBERTa, and GPT dynamically adapt word representations based on their usage, significantly improving the semantic understanding of text.

Generative AI models based on transformers have not only advanced language understanding but also enabled the generation of coherent and contextually relevant text. These models can be leveraged for unsupervised topic discovery by generating topic-related summaries or simulating latent thematic distributions without requiring labeled inputs. When combined with embedding-based similarity measures and clustering algorithms, generative AI facilitates the identification of semantically coherent topics and extraction of representative content, enhancing interpretability and utility.

This combination of generative AI and embedding techniques addresses several limitations of traditional methods. Embedding-based clustering allows flexible topic granularity and adapts better to diverse datasets. Generative models can highlight key sentences or phrases that encapsulate topics, improving human interpretability. Additionally, self-supervised learning paradigms and contrastive learning techniques further enhance the robustness and generalizability of embeddings, making them more resilient to noise and domain shifts.

In this context, unsupervised topic discovery and text extraction using generative AI and embedding techniques have become indispensable tools in various applications, including social media analysis, customer feedback mining, academic research, and legal analytics. This introduction sets the stage for a detailed exploration of these methodologies, their integration, and the advances they bring to unsupervised text mining.

## Literature Survey

Unsupervised topic discovery and text extraction have been extensively studied in natural language processing, evolving significantly over the past decades. Early efforts primarily focused on statistical models such as Latent Dirichlet Allocation (LDA) [Blei et al., 2003], which became a foundational method for uncovering latent topics by modeling documents as mixtures of topics and topics as distributions over words. LDA's probabilistic framework enabled interpretability and scalability but assumed a bag-of-words representation, ignoring word order and deep contextual information. Non-negative Matrix Factorization (NMF) [Lee and Seung, 1999] also gained popularity for topic modeling by decomposing term-document matrices, offering a linear algebraic perspective, but it too faced limitations in capturing semantic richness.

With the emergence of neural networks and representation learning, embedding techniques revolutionized textual analysis. Word2Vec [Mikolov et al., 2013] introduced dense word embeddings that preserved semantic relationships, enabling words with similar meanings to have nearby vector representations. GloVe [Pennington et al., 2014] further improved embeddings by incorporating global co-occurrence statistics. However, both methods generated static embeddings, failing to account for polysemy or context variance.

The introduction of transformer-based models marked a paradigm shift. BERT (Bidirectional Encoder Representations from Transformers) [Devlin et al., 2019] leveraged masked language modeling and deep bidirectional context, producing dynamic embeddings sensitive to word usage. This development enhanced downstream NLP tasks, including unsupervised topic discovery. RoBERTa [Liu et al., 2019] and other transformer variants refined pretraining techniques, boosting embedding quality. Studies such as [Grootendorst, 2020] introduced BERTopic, a framework combining BERT embeddings with clustering algorithms to discover interpretable topics without supervision, illustrating the power of embedding-driven approaches.

Generative AI models, particularly large-scale autoregressive models like GPT series [Radford et al., 2018; Brown et al., 2020], expanded capabilities beyond static representations by generating coherent and contextually relevant text. These models have been adapted for unsupervised tasks, using techniques such as zero-shot and few-shot learning to simulate topic-related content or generate summaries that reveal thematic structures. Research by [Reimers and Gurevych, 2019] demonstrated the effectiveness of sentence embeddings for clustering and semantic search, while recent efforts explore leveraging generative models for extractive summarization [Liu and Lapata, 2019] and keyphrase extraction [Zhang et al., 2020].

The synergy between generative AI and embedding techniques has been further enriched by advancements in self-supervised and contrastive learning. Models like SimCSE [Gao et al., 2021] and SupCon [Khosla et al., 2020] optimize embeddings by contrasting positive and negative pairs, enhancing semantic discrimination even without labels. This improvement translates directly into better clustering and topic extraction performance, addressing challenges related to noisy data and domain adaptation.

Despite these advancements, challenges remain. Topic coherence and interpretability still depend on clustering algorithms and post-processing steps. High computational costs and the need for large-scale pretrained models can limit accessibility. Moreover, generative models occasionally produce plausible but incorrect information, requiring careful validation in critical domains.

Recent surveys [Allahyari et al., 2017; Blei, 2012] and benchmarks underscore the ongoing research in balancing model complexity, interpretability, and scalability. Hybrid approaches that combine statistical methods with neural embeddings and generative models continue to emerge, indicating a trend towards more robust, adaptable frameworks.

In summary, the literature highlights a trajectory from probabilistic topic models to embedding-driven and generative AI-based unsupervised methods, reflecting broader shifts in NLP towards deeper contextual understanding and more flexible, data-efficient solutions. The integration of these technologies offers promising avenues for advancing unsupervised topic discovery and text extraction, with ongoing research focused on overcoming current limitations and expanding practical applications.

## SYSTEM ARCHITECTURE

The system architecture for unsupervised topic discovery and text extraction leveraging generative AI and embedding techniques is designed to efficiently process large volumes of unstructured text data, identify coherent topics, and extract representative content without requiring labeled data. The architecture integrates multiple components, including data ingestion, preprocessing, embedding generation, generative modeling, clustering, and output generation, orchestrated to maximize semantic understanding and interpretability. The following sections detail each component and their interactions within the overall pipeline.

### Data Ingestion and Storage

The system begins with the ingestion of raw textual data, which can come from diverse sources such as social media feeds, news articles, research papers, customer reviews, or internal enterprise documents. This component handles data extraction, transformation, and loading (ETL) to ensure consistent and clean input for downstream processing. Data is stored in a scalable storage layer, such as distributed file systems or cloud object storage, to accommodate potentially large corpora and enable parallel processing.

### Text Preprocessing

Raw textual data often contains noise, inconsistencies, and irrelevant information. The preprocessing module performs several key tasks:

**Tokenization:** Splitting text into tokens (words, phrases, or subwords).

**Normalization:** Lowercasing, removing punctuation, and handling special characters.

**Stop-word removal:** Eliminating common words that do not contribute to semantic meaning.

**Lemmatization/Stemming:** Reducing words to their root forms to unify variants.

**Noise filtering:** Removing irrelevant metadata, advertisements, or HTML tags.

These steps help standardize the input and reduce dimensionality, improving the quality of embeddings and subsequent topic discovery.

**Embedding Generation**

The core of the system is the embedding generation layer, which transforms processed text into dense vector representations capturing semantic and syntactic nuances. This module leverages pretrained transformer-based models such as BERT, RoBERTa, or GPT variants. Embeddings can be generated at multiple granularities:

**Word embeddings:** Capture individual word semantics but may lack context.

**Sentence embeddings:** Represent entire sentences or paragraphs, offering richer context.

**Document embeddings:** Aggregate multiple sentences or entire documents for global understanding.

Contextualized embeddings dynamically adjust representations based on usage, enabling the system to distinguish between polysemous words and complex sentence structures. Depending on the application, the system may also fine-tune these pretrained models on domain-specific corpora to improve relevance.

**Generative AI Module**

This module incorporates generative transformer models that can produce coherent textual content based on learned patterns. Generative AI serves several functions in the architecture:

**Topic summarization:** Generating concise summaries of candidate topics identified through embeddings.

**Keyword/phrase generation:** Proposing representative phrases or key terms that encapsulate topics.

**Topic refinement:** Iteratively improving topic coherence by generating sample documents or sentences under hypothesized topics.

By simulating plausible thematic content, the generative AI assists in enhancing the interpretability and semantic richness of discovered topics. This module operates without supervision, relying on pretrained knowledge and contextual embeddings.

**Clustering and Topic Discovery**

Once embeddings are obtained, the system applies clustering algorithms to group semantically similar documents or sentences into topics. Common clustering techniques include:

**K-means clustering:** Partitioning data into a predefined number of clusters.

**Hierarchical clustering:** Building a tree of clusters to explore topic granularity.

**Density-based clustering (DBSCAN):** Identifying clusters based on data density, useful for noisy datasets.

Distance or similarity metrics, such as cosine similarity, guide the clustering process in the high-dimensional embedding space. The system can automatically estimate the optimal number of topics using heuristics or validation metrics like silhouette scores.

**Topic Representation and Text Extraction**

After clustering, the system extracts representative textual elements for each topic to aid interpretation and usability:

**Representative keywords:** High-frequency or high-weighted terms within clusters.

**Key sentences or excerpts:** Selected via similarity scoring or using generative AI to produce topic summaries.

**Topic labels:** Generated automatically by analyzing dominant keywords or using generative models to propose descriptive labels.

These outputs provide users with clear, concise views of the discovered topics and their core content.

**User Interface and Visualization**

The final layer of the architecture is a user-facing interface that presents the topics and extracted text in an accessible manner. Visualization tools such as interactive topic maps, word clouds, or cluster dendrograms help users explore relationships between topics and drill down into detailed textual content. Additionally, the interface supports search and filtering functionalities to refine results based on user interests.

**Feedback Loop and Model Updating**

To enhance performance over time, the system may incorporate a feedback loop where user interactions and corrections are logged to refine embeddings, generative models, or clustering parameters. This iterative learning process ensures adaptability to evolving datasets and user needs. This system architecture integrates the strengths of embedding techniques and generative AI models within an unsupervised learning framework to achieve scalable, interpretable topic discovery and text extraction. It balances preprocessing rigor, semantic representation, generative capabilities, and clustering robustness to deliver high-quality insights from large unstructured textual corpora. By modularizing components and leveraging pretrained models, the system ensures flexibility and extensibility for diverse applications and domains.

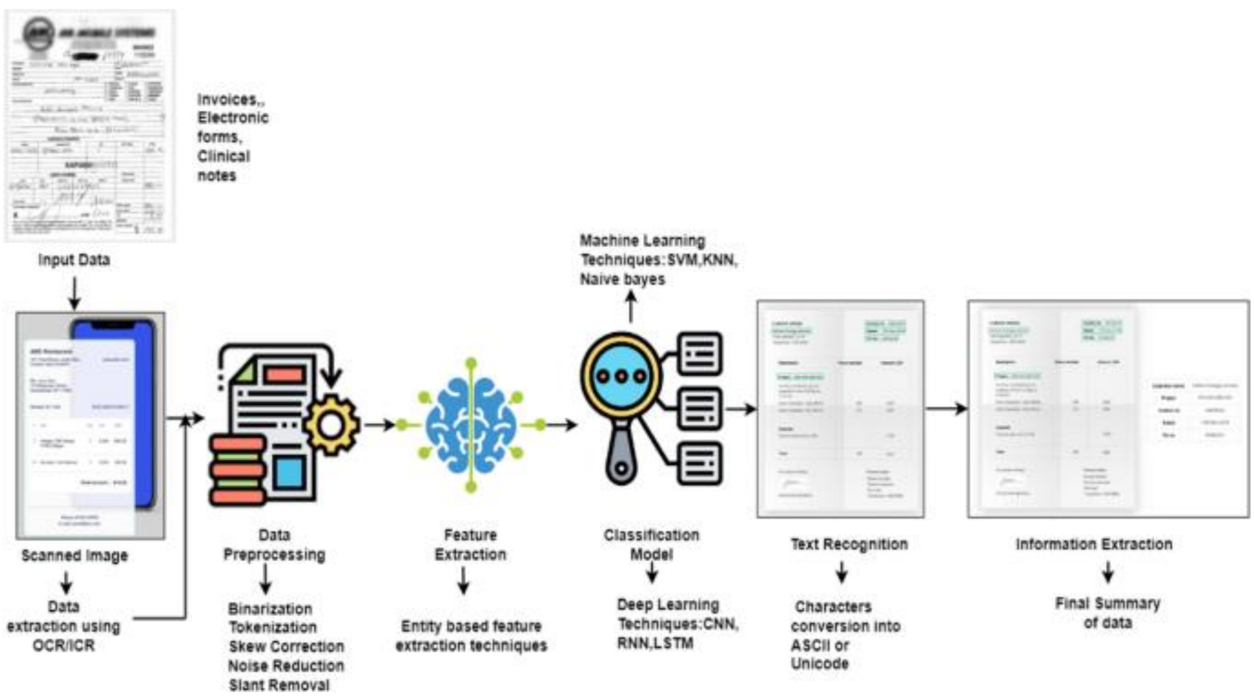


**Fig. 1** Exploring AI-driven approaches for unstructured document analysis and future horizons

## CONCLUSION

The integration of generative AI and embedding techniques marks a significant advancement in the field of unsupervised topic discovery and text extraction. Traditional topic modeling methods, while foundational, often fall short in capturing the rich semantic context and complex language patterns found in modern textual datasets. By leveraging deep contextual embeddings generated by transformer-based models such as BERT and GPT, the system can encode nuanced semantic relationships, enabling more accurate and meaningful topic detection.

Generative AI further complements this by producing coherent summaries and representative content that enhance the interpretability and usability of discovered topics. This synergy allows for flexible clustering and extraction processes that do not rely on labeled data, making it highly applicable to domains where annotated datasets are scarce or costly to obtain. Additionally, advances in self-supervised and contrastive learning have improved the robustness and adaptability of embeddings, addressing issues related to noise and domain shifts.

The proposed system architecture demonstrates how these components can be orchestrated effectively—from data ingestion and preprocessing to embedding generation, generative modeling, clustering, and user-friendly visualization. Such an architecture supports scalable and interpretable analysis of large unstructured corpora, enabling users to gain actionable insights efficiently.

Despite these advances, challenges remain, including computational resource demands, topic coherence evaluation, and ensuring factual accuracy in generative outputs. Future research directions include enhancing model interpretability through explainable AI techniques, integrating multimodal data sources, and developing domain-adaptive models to increase applicability.

Overall, unsupervised topic discovery and text extraction powered by generative AI and embedding methods represent a promising frontier in natural language processing, offering scalable, adaptable, and semantically rich solutions for unlocking insights from vast amounts of textual data.

## REFERENCES:


1. Sidharth, S. (2017). Cybersecurity Approaches for IoT Devices in Smart City Infrastructures.
2. Sidharth, S. (2016). The Role of Artificial Intelligence in Enhancing Automated Threat Hunting 1Mr. Sidharth Sharma.
3. Srinivasan, R. (2025). Friction Stir Additive Manufacturing of AA7075/Al2O3 and Al/MgB2 Composites for Improved Wear and Radiation Resistance in Aerospace Applications. *J. Environ. Nanotechnol*, *14*(1), 295-305.
4. Vijayalakshmi, K., Amuthakkannan, R., Ramachandran, K., & Rajkavin, S. A. (2024). Federated Learning-Based Futuristic Fault Diagnosis and Standardization in Rotating Machinery. *SSRG International Journal of Electronics and Communication Engineering*, *11*(9), 223-236.
5. Sidharth, S. (2019). Enhancing Security of Cloud-Native Microservices with Service Mesh Technologies.
6. Sidharth, S. (2022). Zero Trust Architecture: A Key Component of Modern Cybersecurity Frameworks.
7. Sakthibalan, P., Saravanan, M., Ansal, V., Rajakannu, A., Vijayalakshmi, K., & Vani, K. D. (2023). A Federated Learning Approach for ResourceConstrained IoT Security Monitoring. In *Handbook on Federated Learning* (pp. 131-154). CRC Press.
8. Amuthakkannan, R., & Al Yaqoubi, M. H. A. (2023). Development of IoT based water pollution identification to avoid destruction of aquatic life and to improve the quality of water. *International journal of engineering trends and technology*, *71*(10), 355-370.
9. Sidharth, S. (2016). Establishing Ethical and Accountability Frameworks for Responsible AI Systems.

10. Sidharth, S. (2015). AI-Driven Detection and Mitigation of Misinformation Spread in Generated Content.
11. Amuthakkannan, R., Muthuraj, M., Ademi, E., Rajesh, V., & Ahammad, S. H. (2023). Analysis of fatigue strength on friction stir lap weld AA2198/Ti6Al4V joints. *Materials Today: Proceedings*.
12. Prova, Nuzhat Noor Islam. "Healthcare Fraud Detection Using Machine Learning." *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*. IEEE, 2024.
13. Prova, N. N. I. (2024, August). Garbage Intelligence: Utilizing Vision Transformer for Smart Waste Sorting. In *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)* (pp. 1213-1219). IEEE.
14. Sidharth, S. (2023). AI-Driven Anomaly Detection for Advanced Threat Detection.
15. Sidharth, S. (2023). Homomorphic Encryption: Enabling Secure Cloud Data Processing.
16. Prova, N. N. I. (2024, August). Advanced Machine Learning Techniques for Predictive Analysis of Health Insurance. In *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)* (pp. 1166-1170). IEEE.
17. Prova, N. N. I. (2024, October). Improved Solar Panel Efficiency through Dust Detection Using the InceptionV3 Transfer Learning Model. In *2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)* (pp. 260-268). IEEE.
18. Devi, K., & Indoria, D. (2021). Digital Payment Service In India: A Review On Unified Payment Interface. *Int. J. of Aquatic Science*, *12*(3), 1960-1966.
19. Devi, K., & Indoria, D. (2023). The Critical Analysis on The Impact of Artificial Intelligence on Strategic Financial Management Using Regression Analysis. *Res Militaris*, *13*(2), 7093-7102.
20. Prova, N. N. I. (2024, October). Enhancing Fish Disease Classification in Bangladeshi Aquaculture through Transfer Learning, and LIME Interpretability Techniques. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1157-1163). IEEE.
21. Devi, K., & Indoria, D. (2021). Role of Micro Enterprises in the Socio-Economic Development of Women--A Case Study of Koraput District, Odisha. *Design Engineering*, 1135-1151.
22. Indoria, D. (2021). AN APPLICATION OF FOREIGN DIRECT INVESTMENT. *BIMS International Research Journal of Management and Commerce*, *6*(1), 01-04.
23. Sidharth, S. (2024). Strengthening Cloud Security with AI-Based Intrusion Detection Systems.
24. Sidharth, S. (2022). Enhancing Generative AI Models for Secure and Private Data Synthesis.
25. Kumar, G. H., Raja, D. K., Suresh, S., Kottamala, R., & Harsith, M. (2024, August). Vision-Guided Pick and Place Systems Using Raspberry Pi and YOLO. In *2024 2nd International Conference on Networking, Embedded and Wireless Systems (ICNEWS)* (pp. 1-7). IEEE.
26. Kumar, G. H., Raja, D. K., Varun, H. D., & Nandikol, S. (2024, November). Optimizing Spatial Efficiency Through Velocity-Responsive Controller in Vehicle Platooning. In *2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS)* (pp. 1-5). IEEE.
27. Sidharth, S. (2021). Multi-Cloud Environments: Reducing Security Risks in Distributed Architectures.
28. Sidharth, S. (2020). The Rising Threat of Deepfakes: Security and Privacy Implications.
29. Kalimuthu, S., Perumal, T., Yaakob, R., Marlisah, E., & Babangida, L. (2021, March). Human Activity Recognition based on smart home environment and their applications, challenges. In *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 815-819). IEEE.

30. Vidhyasagar, B. S., Lakshmanan, A. S., Abishek, M. K., & Kalimuthu, S. (2023, October). Video captioning based on sign language using yolov8 model. In *IFIP International Internet of Things Conference* (pp. 306-315). Cham: Springer Nature Switzerland.
31. Thamma, S. R. T. S. R. (2024). Optimization of Generative AI Costs in Multi-Agent and Multi-Cloud Systems.
32. Thamma, S. R. T. S. R. (2024). Revolutionizing Healthcare: Spatial Computing Meets Generative AI.
33. Vidhyasagar, B. S., Arvindhan, M., Arulprakash, A., Kannan, B. B., & Kalimuthu, S. (2023, November). The crucial function that clouds access security brokers play in ensuring the safety of cloud computing. In *2023 International Conference on Communication, Security and Artificial Intelligence (ICCSAI)* (pp. 98-102). IEEE.
34. Vidhyasagar, B. S., Harshagnan, K., Diviya, M., & Kalimuthu, S. (2023, October). Prediction of Tomato Leaf Disease Plying Transfer Learning Models. In *IFIP International Internet of Things Conference* (pp. 293-305). Cham: Springer Nature Switzerland.
35. Thamma, S. R. (2024). Cardiovascular image analysis: AI can analyze heart images to assess cardiovascular health and identify potential risks.
36. Thamma, S. R. T. S. R. (2024). Generative AI in Graph-Based Spatial Computing: Techniques and Use Cases.
37. Kalimuthu, S., Perumal, T., Yaakob, R., Marlisah, E., & Raghavan, S. (2024, March). Multiple human activity recognition using iot sensors and machine learning in device-free environment: Feature extraction, classification, and challenges: A comprehensive review. In *AIP Conference Proceedings* (Vol. 2816, No. 1). AIP Publishing.
38. Bs, V., Madamanchi, S. C., & Kalimuthu, S. (2024, February). Early Detection of Down Syndrome Through Ultrasound Imaging Using Deep Learning Strategies—A Review. In *2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE)* (pp. 1-6). IEEE.
39. Turlapati, V. R., Vichitra, P., Raval, N., Khaja Mohinuddeen, J., & Mishra, B. R. (2024). Ethical Implications of Artificial Intelligence in Business Decision-making: A Framework for Responsible AI Adoption. *Journal of Informatics Education and Research*, *4*(1).
40. Raju, P., Arun, R., Turlapati, V. R., Veeran, L., & Rajesh, S. (2024). Next-Generation Management on Exploring AI-Driven Decision Support in Business. In *Optimizing Intelligent Systems for Cross-Industry Application* (pp. 61-78). IGI Global.
41. Kalimuthu, S., Perumal, T., Marlisah, E., Yaakob, R., BS, V., & Ismail, N. H. (2024). HUMAN ACTIVITY RECOGNITION BASED ON DEVICE-FREE WI-FI SENSING: A COMPREHENSIVE REVIEW. *Malaysian Journal of Computer Science*, *37*(3), 252-269.
42. Seshanna, M., Kumar, H., Seshanna, S., & Alur, N. (2021). THE INFLUENCE OF FINANCIAL LITERACY ON COLLECTIBLES AS AN ALTERNATIVE INVESTMENT AVENUE: EFFECTS OF FINANCIAL SKILL, FINANCIAL BEHAVIOUR AND PERCEIVED KNOWLEDGE ON INVESTORS'FINANCIAL WELLBEING. *Turkish Online Journal of Qualitative Inquiry*, *12*(4).
43. Rao, P. S. (2008). *International Business Environment*. HIMALAYA PUBLISHING HOUSE 2nd Rev. ed..
44. Deshmukh, M., Ghadle, K., & Jadhav, O. (2020). An innovative approach for ranking hexagonal fuzzy numbers to solve linear programming problems. *International Journal on Emerging Technologies*, *11*(2), 385-388.

45. Patil, R. D., & Jadhav, O. S. (2016). Some contribution of statistical techniques in big data: a review. *International Journal on Recent and Innovation Trends in Computing and Communication*, *4*(4), 293-303.
46. Sreekanthaswamy, N., Anitha, S., Singh, A., Jayadeva, S. M., Gupta, S., Manjunath, T. C., & Selvakumar, P. (2025). Digital Tools and Methods. *Enhancing School Counseling With Technology and Case Studies*, *25*.
47. Sreekanthaswamy, N., & Hubballi, R. B. (2024). Innovative Approaches To Fmcg Customer Journey Mapping: The Role Of Block Chain And Artificial Intelligence In Analyzing Consumer Behavior And Decision-Making. *Library of Progress-Library Science, Information Technology & Computer*, *44*(3).
48. Kalluri, V. S. Impact of AI-Driven CRM on Customer Relationship Management and Business Growth in the Manufacturing Sector. *International Journal of Innovative Science and Research Technology (IJISRT)*.
49. Kalluri, V. S. Optimizing Supply Chain Management in Boiler Manufacturing through AI-enhanced CRM and ERP Integration. *International Journal of Innovative Science and Research Technology (IJISRT)*.
50. Nair, S. S., Lakshmikanthan, G., Kendyala, S. H., & Dhaduvai, V. S. (2024, October). Safeguarding Tomorrow-Fortifying Child Safety in Digital Landscape. In *2024 International Conference on Computing, Sciences and Communications (ICCSC)* (pp. 1-6). IEEE.
51. Lakshmikanthan, G., Nair, S. S., Sarathy, J. P., Singh, S., Santiago, S., & Jegajothi, B. (2024, December). Mitigating IoT Botnet Attacks: Machine Learning Techniques for Securing Connected Devices. In *2024 International Conference on Emerging Research in Computational Science (ICERCS)* (pp. 1-6). IEEE.
52. Kalluri, S. V. S., & Narra, S. (2024). Predictive Analytics in ADAS Development: Leveraging CRM Data for Customer-Centric Innovations in Car Manufacturing. *vol*, *9*, 6.
53. Kalluri, V. S., Malineni, S. C., Seenivasan, M., Sakkarai, J., Kumar, D., & Ananthan, B. (2025). Enhancing manufacturing efficiency: leveraging CRM data with Lean-based DL approach for early failure detection. *Bulletin of Electrical Engineering and Informatics*, *14*(3), 2319-2329.
54. Nair, S. S. (2023). Digital Warfare: Cybersecurity Implications of the Russia-Ukraine Conflict. *International Journal of Emerging Trends in Computer Science and Information Technology*, *4*(4), 31-40.
55. Chu, T. S., Nair, S. S., & Lakshmikanthan, G. (2022). Network Intrusion Detection Using Advanced AI Models A Comparative Study of Machine Learning and Deep Learning Approaches. *International Journal of Communication Networks and Information Security (IJCNIS)*, *14*(2), 359-365.
56. Jeyaprabha, B., & Sundar, C. (2021). The mediating effect of e-satisfaction on e-service quality and e-loyalty link in securities brokerage industry. *Revista Geintec-gestao Inovacao E Tecnologias*, *11*(2), 931-940.
57. Jeyaprabha, B., & Sunder, C. What Influences Online Stock Traders’ Online Loyalty Intention? The Moderating Role of Website Familiarity. *Journal of Tianjin University Science and Technology*.
58. Lakshmikanthan, G., & Nair, S. S. (2024). Protecting Self-Driving Vehicles from attack threats. *International Journal of Emerging Research in Engineering and Technology*, *5*(1), 16-20.
59. Sivakumar, K., Manoj Kumar, S., Saravanan, G., & Mahendran, G. (2025). Mechanical, wear, fatigue, water absorption and flammability of silane-treated Indian squid chitin powder-dispersed pineapple fiber-polyester composite. *Polymer Bulletin*, *82*(5), 1663-1683.

60. Mahendran, G., Kumar, S. M., Uvaraja, V. C., & Anand, H. (2025). Effect of wheat husk biogenic ceramic Si3N4 addition on mechanical, wear and flammability behaviour of castor sheath fibre-reinforced epoxy composite. *Journal of the Australian Ceramic Society*, 1-10.
61. Jeyaprabha, B., Catherine, S., & Vijayakumar, M. (2024). Unveiling the Economic Tapestry: Statistical Insights Into India's Thriving Travel and Tourism Sector. In *Managing Tourism and Hospitality Sectors for Sustainable Global Transformation* (pp. 249-259). IGI Global.
62. JEYAPRABHA, B., & SUNDAR, C. (2022). The Psychological Dimensions Of Stock Trader Satisfaction With The E-Broking Service Provider. *Journal of Positive School Psychology*, 3787-3795.
63. Mahendran, G., Mageswari, M., Kakaravada, I., & Rao, P. K. V. (2024). Characterization of polyester composite developed using silane-treated rubber seed cellulose toughened acrylonitrile butadiene styrene honey comb core and sunn hemp fiber. *Polymer Bulletin*, *81*(17), 15955-15973.
64. Mahendran, G., Gift, M. M., Kakaravada, I., & Raja, V. L. (2024). Load bearing investigations on lightweight rubber seed husk cellulose–ABS 3D-printed core and sunn hemp fiber-polyester composite skin building material. Macromolecular Research, 32(10), 947-958.
65. Kumar, T. V. (2019). Cloud-Based Core Banking Systems Using Microservices Architecture.
66. Kumar, T. V. (2019). BLOCKCHAIN-INTEGRATED PAYMENT GATEWAYS FOR SECURE DIGITAL BANKING.
67. Mohanavel, V., Diwakar, G., Govindasamy, M., Singh, V., Theophilus Rajakumar, I. P., Soudagar, M. E. M., ... & Alharbi, S. A. (2024). Fabrication of ramie/hemp fibers-reinforced hybrid polymer composite—A comprehensive study on biological and structural application. *AIP Advances*, *14*(8).
68. Nadaf, A. B., Sharma, S., & Trivedi, K. K. (2024). CONTEMPORARY SOCIAL MEDIA AND IOT BASED PANDEMIC CONTROL: A ANALYTICAL APPROACH. *Weser Books*, 73.
69. Trivedi, K. K. (2022). A Framework of Legal Education towards Litigation-Free India. *Issue 3 Indian JL & Legal Rsch.*, *4*, 1.
70. Kumar, T. V. (2015). CLOUD-NATIVE MODEL DEPLOYMENT FOR FINANCIAL APPLICATIONS.
71. Kumar, T. V. (2018). REAL-TIME COMPLIANCE MONITORING IN BANKING OPERATIONS USING AI.
72. Trivedi, K. K. (2022). HISTORICAL AND CONCEPTUAL DEVELOPMENT OF PARLIAMENTARY PRIVILEGES IN INDIA.
73. Himanshu Gupta, H. G., & Trivedi, K. K. (2017). International water clashes and India (a study of Indian river-water treaties with Bangladesh and Pakistan).
74. Trivedi, K. K. (2017). Cultural Influences On The Effectiveness Of Women Protection Legislation.
75. Kumar, T. V. (2020). Generative AI Applications in Customizing User Experiences in Banking Apps.
76. Kumar, T. V. (2020). FEDERATED LEARNING TECHNIQUES FOR SECURE AI MODEL TRAINING IN FINTECH.
77. Hussain, M. I., Shamim, M., Ravi Sankar, A. V., Kumar, M., Samanta, K., & Sakhare, D. T. (2022). The effect of the Artificial Intelligence on learning quality & practices in higher education. *Journal of Positive School Psychology*, 1002-1009.
78. Prasad, V., Dangi, A. K., Tripathi, R., & Kumar, N. (2023). Educational Perspective of Intellectual Property Rights. *Russian Law Journal*, *11*(2S), 257-268.
79. Kumar, T. V. (2022). AI-Powered Fraud Detection in Real-Time Financial Transactions.
80. Kumar, T. V. (2021). NATURAL LANGUAGE UNDERSTANDING MODELS FOR PERSONALIZED FINANCIAL SERVICES.

81. Khachariya, H. D., Naveen, S., Al-Nussairi, A. K. J., Abood, B. S. Z., Alanssari, A. I., & Shaker, Z. Y. (2024, November). Deep Learning for Workforce Planning and Analytics. In *2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES)* (pp. 1-5). IEEE.