Secure Edge AI for Financial Document Processing

Olivia Johnson
Asst Professor

Abstract

Enhancing security and privacy during feature extraction of financial documents is of paramount importance. Conventional concerns of data loss during feature extraction can be addressed in a natural way by placing feature extraction on the edge layer of an edge-cloud ecosystem. Recent developments in Generative and Agentic AI open opportunities for generating features for various tasks without sharing or leaking sensitive details about the data. Such capabilities can be expressed as Generative and Agentic AI-enhanced feature extraction—combining agentic capabilities of Generative AI to supervise the feature extraction process and also provide additional semantics and or representations that help improve the accuracy of the downstream task. This framework is shown to naturally extend to both edge-layer and cloud-layer processing. Security and privacy requirements during the feature-extraction stage are met by suitably applying standard security mechanisms without any changes to the underlying principle or operation of the feature extraction process.

Generative AI is an emerging technology with numerous applications in various domains. The recent capability of generating new data and labels from a few or a small set of samples but without sharing or leaking the private details in the original data has opened another important dimension. This capability not only allows the origin of the samples to remain secret, but also helps protect the sensitive nature of the data. Generative AI can thus accentuate the benefits of decentralization by further reducing the role and dominance of central-proxy-service providers in edge and cloud ecosystems. Instead of consuming and storing all the edge- and cloud-layer data, the major job can be reduced to merely providing the infrastructure for empowering other users in the ecosystem with the appropriate private-label Generation AI capability required for their workloads.

Keywords: Edge Cloud Ecosystem, Secure Feature Extraction, Financial Document Privacy, Generative AI Feature Synthesis, Agentic AI Supervision, Privacy Preserving Machine Learning, Edge Layer Intelligence, Cloud Layer Analytics, Decentralized AI Architecture, Sensitive Data Protection, Semantic Feature Enrichment, Downstream Task Accuracy, Synthetic Feature Generation, Private Label AI Models, Zero Data Leakage Design, Secure AI Pipelines, Confidential Document Processing, Trustworthy AI Systems, Edge AI Security, Federated Intelligence.

1. Introduction

Each day, business enterprises scrutinize and process a considerable number of financial documents that require feature extraction for advanced processing. Mining highly sensitive data via Optical Character Recognition (OCR) constitutes one of the prime steps in feature extraction of these documents. However, the OCR step could be a prime target for confidential information leakage due to the huge corpus of documents that are processed daily (creating a prime space for adversarial attacks) or due to the fact that these documents are processed externally in a cloud service provider (leading to potential concern for security and privacy of the semantically rich extracted features).
The emergence of Generative and Agentic AI technologies serves as an opportunity for parallelizing the generation of the BERT embeddings and for minimizing or avoiding the use of OCR by generating common features in a secure way. Therefore, design and development of a hybrid edge–cloud architecture for financial document processing that

ISSN: 3067-4190
JOURNAL OF ARTIFICIAL INTELLIGENCE AND BIG DATA DISCIPLINES
Volume: 03 Issue: 03
RECEIVED: JULY 22 REVISED: AUGUST 06 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 04

leverages Generative and Agentic AI technologies for improving the execution speed and security of the feature-extraction step is suggested here. The proposed architecture has an emphasis on the security, privacy, and exploitation of the Generative and Agentic AI concepts in edge–cloud ecosystems, but it can be extended to any computation in an edge–cloud ecosystem, independently of the AI topic.

1.1. Overview of Generative and Agentic AI Concepts

Generative and agentic artificial intelligence are maturing areas of research and development, with major impacts on society. Generative AI capabilities form part of many of the latest products and services, while agentic AI promises to extend and automate the use of generative capabilities, among others. A detailed enablement model is described for these two areas. The framework identifies concepts, ideas, essential components, and interactions of both areas, allowing for more informed academic and business discussions, product understanding and development, and implication and impact analysis.

Generative and agentic artificial intelligence (AI) are maturing areas of research and development, with major impacts on society. Generative AI capabilities are embedded in many of the latest products and services, while agentic AI is emerging to extend and automate the use of generative capabilities, among others. A detailed enablement model captures these areas. The framework identifies concepts, ideas, essential components, and interactions of both areas, allowing for more informed academic and business discussions, product understanding and development, and implication and impact analysis.

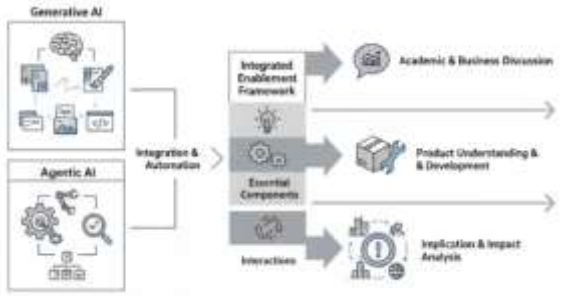


Fig 1: Synthesizing Generative and Agentic Intelligence: A Unified Enablement Framework for Socio-Technical Impact and Governance

2. Background and Motivation

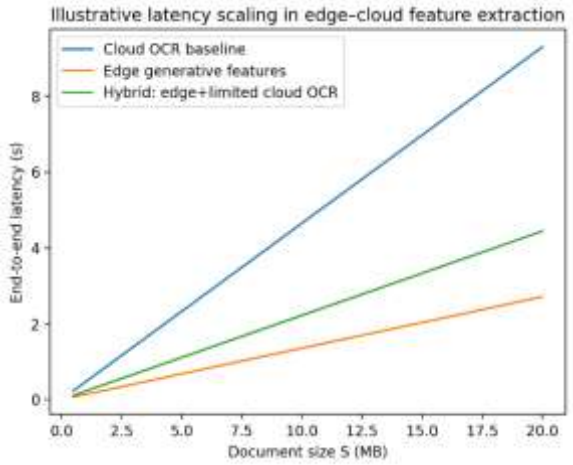
The concepts of generative AI and agentic AI are experiencing explosive growth, propelled by recent technical and theoretical advances in AI. The technological implications of these approaches are tractable and several important and overarching social, governance, and legal questions are being addressed. A Framework for Understanding Generative and Agentic AI indicates the significant and potentially damaging impact of current generative AI tools and suggests a perspective as both tool and medium for agentic AI. The positive and negative implications for data-intensive research in general and for financial documents specifically are still emerging. Artificial intelligence-based feature extraction transformations from documents stored at the network edge to network core are critically dependent on privacy and security considerations. Despite the rising number of data breaches targeting cloud services, the cloud remains attractive because of the massive computational power and advanced AI services available for processing large amounts of data. Financial organisations continue to leverage cloud computing services to deploy various services, including core banking and information services. Generative AI has fascinated the public with artificial image, video, and music generations. However, users also express concerns about when, how, and why data are collected and used, especially when their privacy is compromised.

2.1. Impacts and Implications of Generative and Agentic AI

Generative and agentic artificial intelligence (AI) is revolutionizing many industrial and governmental processes while introducing new risks in cybersecurity, destabilizing employment, enabling the generation of realistic fake news, and facilitating the disinformation campaigns of governments as well as organized crime

ISSN: 3067-4190
JOURNAL OF ARTIFICIAL INTELLIGENCE AND BIG DATA DISCIPLINES
Volume: 03 Issue: 03
RECEIVED: JULY 22 REVISED: AUGUST 06 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 04

groups. Generative AI algorithms can produce videos, news articles, images, software codes, and 3-D visualizations that seem indistinguishable from those created by humans. However, these products typically lack the creativity, common sense, and critical thinking of human artisans. Agentic AI is an embodiment of artificial intelligence in autonomous physical agents such as intelligent robots and self-driving cars, among others. In these use cases, generative models can provide sensory inputs as part of the agentic system, enhancing its perception and creating virtual environments that facilitate reinforcement learning. A plethora of generative AI products are flooding the market, and managed purchases are changing everything from marketing and product development to trip planning and software development. Thus, many companies are hooking up generative applications to their vast repositories of proprietary data in an attempt to achieve a competitive edge. This novel framework helps integrate generative systems in a way that ensures the confidentiality and proper usage of sensitive data and models while eliminating adversarial samples through abstaining, steering, and filtering mechanisms. The considered generative models are useful for private operations involving cognitively demanding tasks, such as financial document feature extraction with a possibility of use across edge and cloud layers.



Equation 1: Formalizing the edge–cloud feature-extraction pipeline

1.1 Variables (document → features)

Let:

- Raw document (image/PDF) be x
- Extracted/synthesized feature vector be $z \in \mathbb{R}^d$
- Downstream label/task output be y (fraud flag, KYC class, default risk, etc.)
- Feature extractor be $F(\cdot)$, classifier be $C(\cdot)$

Baseline (OCR-heavy, cloud):

$$z = F_{\text{ocr}}(x), \quad \hat{y} = C(z)$$

So model it as:

$s = G_{\text{desc}}(x)$ (generative description/summary) $z = F_{\text{feat}}(s)$ (feature extraction from description) $\hat{y} = C(z)$

3. Theoretical Foundations of Generative and Agentic AI

Generative artificial intelligence (AI) enables a machine to create or generate new contents by training on the underlying distribution of a set of training samples. Conversely, agentic AI endows a machine with self-awareness, inspector functions and the ability to become an agent that accepts tasks and delegate them to AI tools. Recent years also observed a growing interest in the study of the implications of both kinds of AI models on society, the economy, science, the environment, security, and the applicability of theories, among others. Three viewpoints dominate. First, current AI systems lack general intelligence, consciousness, and the ability to independently improve their technologies and findings. Second, generative deep-learning engines draw statistically plausible conclusions but lack understanding and generalization capabilities. Third, AI systems do not have the ability to use those generated contents in the real scenarios. Examination of scientific, economic, and technologic history reveals that technological innovation

ISSN: 3067-4190
JOURNAL OF ARTIFICIAL INTELLIGENCE AND BIG DATA DISCIPLINES
Volume: 03 Issue: 03
RECEIVED: JULY 22 REVISED: AUGUST 06 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 04

has been driven by the emergence and use of superior intelligent tools.

Generative AI can be described as the development of Holistic Intelligence-1 tools, capable of integrating and exploiting available Holistic Intelligence-1 solutions. Holistic Intelligence-1 is the combined intelligence generated by the use of all available intelligent machines and tools to address complex missions, especially in scientific discovery and business. Scientific research on agentic AI emphasizes its potential applications and implications, including the formulation of complex tasks and flowcharts for their sequential performance, integration and improvement of generative AI products, creation of AI detective tools and privacy agents, and reinforcement of human authors regarding ethics and the achievement of major goals. The intelligent tools that would generate Holistic Intelligence-2, including consciousness, have not yet been developed.

Fig 2: From Generative Synthesis to Agentic Governance: A Framework for Holistic Intelligence-1 and the Evolution of Autonomous Technical Systems

3.1. Conceptual Framework for Understanding Generative and Agentic AI

Although the generative and agentic perspectives on AI systems mainly relate to their usability and programming, they also provide valuable insights into their impacts on the economy and society, as well as into the theory and foundations of AI. A formal definition of the concept of generative AI is still absent from both scholarly literature and authoritative sources; however, it can be viewed as a description of specific capabilities and properties of AI systems that are either central to the AI application domain or that make such systems exploitable and valuable to end users. Consequently, a definition of agentic AI can be offered indirectly. Such a definition would be built upon discussions of generative AI and of other AI capabilities; it would focus on the usability of AI systems and on the programming experience they offer. Such a conceptual framework contributes to understanding the relationship between generative and agentic AI and to recognising which AI properties are primarily relevant for system impacts.

The economic, social, political, and scientific implications of generative and agentic AI transcend the safety and ethical concerns traditionally associated with AI technologies. The paradigm is indeed so disruptive that early warnings have emerged from various fields, including machine learning, software engineering, politics, and social development, as well as cultural awareness. However, such reports often neglect the importance of balancing risks and opportunities to ensure societal advancement. Central to this complex and important process is the correct identification of the potential benefits and threats posed by generative and agentic AI; consideration of these dimensions can clarify what should be feared, what should be safeguarded, or what should be controlled to avoid economic and social instability.

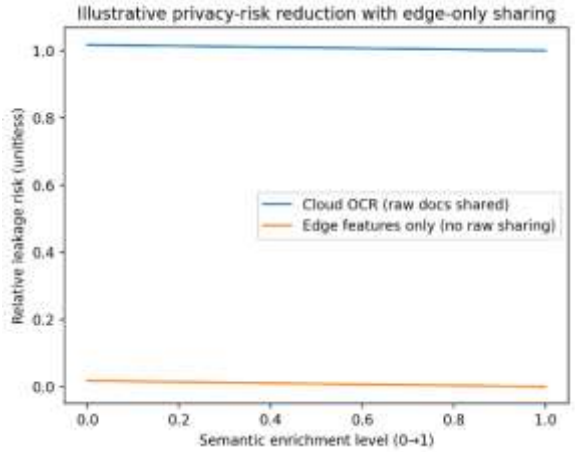
4. Framework for Edge-Cloud Feature Extraction

A conceptual framework illustrates how feature extraction can employ generative and agentic AI to meet security and privacy requirements for sensitive documents. Financial documents are particularly valuable on the dark web, and securing feature extraction for a Financial Technology company is challenging because the data processing takes place at the edge in a multi-user setting, while the feature extraction responsibility resides with the cloud. Part of the solution includes a generative model that jointly learns representations for short-term loan contracts and account statements.

An agentic AI system binds together existing autonomous AI tools. The architecture enables the recognition of features on financial documents useful for fraud detection,

ISSN: 3067-4190
JOURNAL OF ARTIFICIAL INTELLIGENCE AND BIG DATA DISCIPLINES
Volume: 03 Issue: 03
RECEIVED: JULY 22 REVISED: AUGUST 06 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 04

such as identity theft or loan scams, as well as predicting the likelihood of a loan default. To enhance the recognition quality without user intervention, the system uses generative models for exploitation. The system architecture also incorporates security mechanisms that safeguard the data and lessen the chance of leakage of the learned features.User data, including the documents, are encrypted before transmission. Only the keys for the user data that belong to the feature extractor agent are made available to the feature-extraction module in the cloud.



Equation 2: Bimodal (visual + textual) feature extraction equations

2.1 Two encoders (visual/text) → latent codes

Let:

- visual input x_v (document image/regions)
- textual input x_t (OCR text, or generated description s, or both)

Encoders:

$$h_v = E_v(x_v) \in \mathbb{R}^{d_v}, \qquad h_t = E_t(x_t) \in \mathbb{R}^{d_t}$$

2.2 Fusion into a joint representation

Common fusion choices (any of these fits the paper's description):

(A) Concatenation + projection

$$z = W\,[h_v; h_t] + b$$

(B) Gated fusion

$$g = \sigma\big(W_g[h_v; h_t] + b_g\big)\; z = g \odot h_v + (1-g) \odot h_t$$

(C) Attention-based fusion (document sections)
If the document is split into n sections $x^{(i)}$:

$$h^{(i)} = \phi\big(x^{(i)}\big), \quad \alpha_i = \frac{e^{u^\top h^{(i)}}}{\sum_{j=1}^{n} e^{u^\top h^{(j)}}}, \quad z = \sum_{i=1}^{n} \alpha_i\, h^{(i)}$$

4.1. Architectural Overview

Generative and agentic artificial intelligence (AI) have the potential to deliver innovations across various domains, from economics and education to health and technology. The interaction of these technologies could reshape society in many ways, some of them currently unforeseeable. To realise their potential, however, the technologies must be controlled carefully.

The proposed framework supports automated feature extraction of financial documents, expanding the privacy-preserving capabilities of both centralised and decentralised AI approaches. Using a generative AI model at the edge collects labelled features from a financial document dataset. These features are then stored in an encryption protocol that guarantees access by the remotely executed agentic AI—without exposing any sensitive information on the financial document itself. The edge-cloud design is suitable for both centralised and decentralised AI paradigms. Agent-based models offer a way to involve and simulate the behaviour, motivation, and interaction of autonomous entities, while integrating the generative capabilities of deep learning advances presents a systematic way to obtain an optimised training dataset for specialised tasks in environments with unsupervised contexts. The framework considers the extraction of features related specifically to

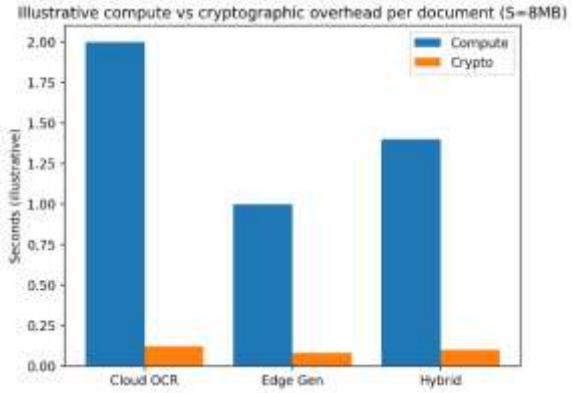
ISSN: 3067-4190
JOURNAL OF ARTIFICIAL INTELLIGENCE AND BIG DATA DISCIPLINES
Volume: 03 Issue: 03
RECEIVED: JULY 22 REVISED: AUGUST 06 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 04

financial documents but can be applied to other areas as well.

4.2. Data Flow and Privacy-Preserving Mechanisms

Serialized images of documents and structured representations of their features seamlessly flow across an edge-cloud ecosystem to extract their most applicable content—semantic, natural-language textual descriptions of the documents or their individual sections. The document-feature-generating generative model, hosted in the cloud, receives the serialized images of the documents as its sole input. The process employs internally stored knowledge to produce a set of features that allow agents to understand at least partially the perceived documents, label them semantically, or even generate an entire document in natural language describing the contents of the perceived document. After completing the description-generation task, the edge-hosted agent deletes the context coma memory to ensure the privacy of any sensitive information of the perceived document.

The feature-extraction step exploits textual descriptions of the documents' contents instead of the documents themselves, and each document is encrypted and can be accessed only by the agent authorized to encrypt it. The security reinforced by addressing the potential threats at a data-flow level is complemented by the privacy-preserving feature-extraction mechanisms that adopt the concept of access control from the twist of using encryption to allow access only to authorized users.



5. Security and Privacy Considerations

Security, privacy, and confidence in fully automated feature extraction results remain major obstacles toward practical real-world solutions. Separation of data sources and processing is also required due to regulations surrounding finance. The architectural design presented in this work mitigates security risks, implements additional privacy-preserving mechanisms, and models the underlying threats posed by malicious actors who gain unauthorized access to either the cloud or edge environment. Based on this model, five countermeasures are proposed to defend against data containment or modification attacks. The client generates a semantically-enriched representation of financial documents to restrict feature extraction model access to the minimum number of supported labels. The documents are then encrypted to prevent their contents from being revealed to the feature extraction services at the edge cloud. The labels of the semantically-enriched representations that the feature extraction model is allowed to access are encrypted with Literal Encryption for privacy-preserving access control. The classification layer on the feature extraction model is trained within a cryptographic secret-sharing setting to enable inference over encrypted data.

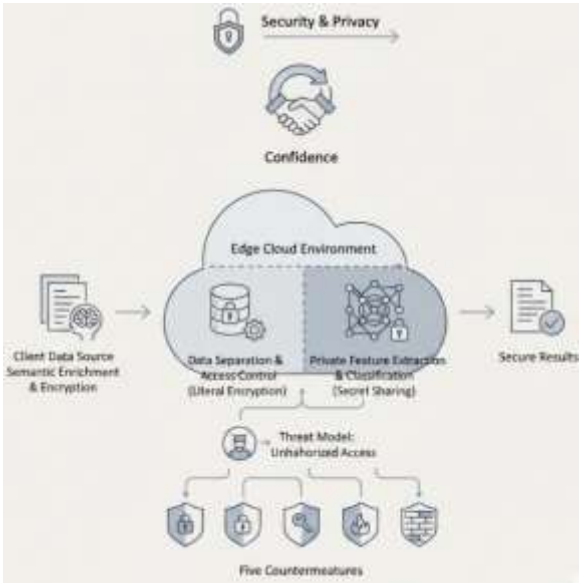


Fig 3: Privacy-Preserving Automated Feature Extraction: A Multi-Layered Cryptographic Architecture for Secure Financial Document Processing

5.1. Threat Model

Document features are sensitive and follow a confidential lifecycle, such as the original document, schema, entity mapping, dataset, and loaded database table structure. Processes appearing at layer i of the edge-cloud architecture must be authorized to access these features. To prevent data sharing with unauthorized processes, data encryption can be performed. For example, data stored in the cloud or sent over the network can be encrypted using feature-specific symmetric keys generated using a key derivation function. A data leakage scenario can result from exploitation of a vulnerability in a process that operates at layer i. For example, the enterprise control centre requires access to manually edited empirical data, which, when misused, can affect the reliability of the deployed quantitative method. The leakage of the knowledge base of document level classification can also have a serious impact, but it is important to note that a semantically similar but syntactically different classification cannot be recognised by any document classifier. Agent-based anomaly detection can be considered for detecting misuse of this type of highly sensitive information.

The framework supports other complementary techniques that can be applied to any part of the document feature extraction process across all three layers of the edge-cloud ecosystem. First, deep learning models (e.g., for spam detection, natural language processing, image classification) can be used to build, pre-train, or fine-tune deep learning classifiers. These models do not need complicated or time-consuming labelling and can be generated directly from the data patterns. Such data patterns can also automatically generate test cases for supervised classification or detection. Moreover, deep generation models (e.g., for generative design, image generation) can use two-level generative meta-modeling, where the high-level generative meta-model synthesizes a hierarchical generative meta-model and synthesizes data patterns for low-level generation processes. The capabilities of agent-based computational discovery can also support procedures at multiple abstraction levels.

Symbol	Meaning	Typical low	Typical high	Notes
S	Document size (MB)	0.5	20.0	Input PDF/image payload
r_{ocr}	OCR throughput (MB/s)	1.0	10.0	Edge or cloud OCR speed
r_{gen}	Generative feature synthesis rate (MB/s)	2.0	20.0	Edge model throughput
$B_{\uparrow}$	Uplink bandwidth (MB/s)	0.5	50.0	Edge → cloud
$B_{\downarrow}$	Downlink bandwidth (MB/s)	0.5	50.0	Cloud → edge
o_{enc}	Encryption overhead fraction	0.02	0.20	Payload expansion (IV/MAC/headers)
γ	Feature size fraction of raw doc	0.01	0.10	Features are smaller than documents
ρ	Fraction of doc sent for OCR (hybrid)	0.05	0.50	Selected regions only
d	Embedding dimension	64	1024	Size of feature vector

ISSN: 3067-4190
JOURNAL OF ARTIFICIAL INTELLIGENCE AND BIG DATA DISCIPLINES
Volume: 03 Issue: 03
RECEIVED: JULY 22 REVISED: AUGUST 06 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 04

Symbol	Meaning	Typical low	Typical high	Notes
m	# of feature labels	5	200	Policy-managed semantic labels

5.2. Encryption and Access Control

Although financial documents are usually stored in a cloud service (offering large storage space as well as low costs), the processing often requires the joint effort of multiple involved parties (e.g., the bank and the client). Their joint access to semi-structured information is usually provided through a digital channel or a Web service. However, this brings the risk of exposing sensitive data to several parties. Therefore, traffic encryption and fine-grained access control might not be sufficient to fully protect the content of the data exchanged and stored. On the one hand, the document itself can be decrypted only by the party to whom it has been encrypted or for which it has been made accessible (to keep the addressed parties content-hidden). On the other hand, the information structure may also be hidden to parties other than the ones who actually need access to it (to keep the addressed parties context-hidden). These issues can be addressed through the application of cryptographic data hiding: (i) data that must remain confidential needs to be encrypted, taking advantage of symmetric encryption (for low-cost operations and good performance) or public key encryption (to benefit from fine-grained permission assignment) schemes; (ii) semi-structured data and documents need to be concealed via a semantic-preserving data-hiding approach that encrypts concepts and label values into oblivion ciphers, while still enabling queries and operations over the document features; and (iii) traffic should be additionally mixed through a cover mechanism on top of encryption, to resist adversaries capable of performing traffic analysis (e.g., pinch-and-zoom, pattern recognition).

Equation 3: GAN objective used for "feature synthesis" (step-by-step)

3.1 Discriminator's role

$D(\cdot)$ outputs probability "real".

- Real samples: $z \sim p_{\text{real}}(z)$ (true features)
- Fake samples: $\tilde{z} = G(\epsilon)$, $\epsilon \sim p(\epsilon)$ noise (or conditional input)

So:

$$D(z) = P(\text{real} \mid z)$$

3.2 Discriminator loss (maximize correct classification)

Maximize:

$$\mathbb{E}_{z\sim p_{\text{real}}}[\log D(z)] + \mathbb{E}_{\epsilon\sim p(\epsilon)}\left[\log\left(1 - D\big(G(\epsilon)\big)\right)\right]$$

Equivalently, define the **minimization** loss:

$$\mathcal{L}_D = -\mathbb{E}_{z\sim p_{\text{real}}}[\log D(z)] - \mathbb{E}_{\epsilon\sim p(\epsilon)}\left[\log\left(1 - D\big(G(\epsilon)\big)\right)\right]$$

3.3 Generator loss (fool the discriminator)

Generator wants $D\big(G(\epsilon)\big) \to 1$. Common "non-saturating" form:

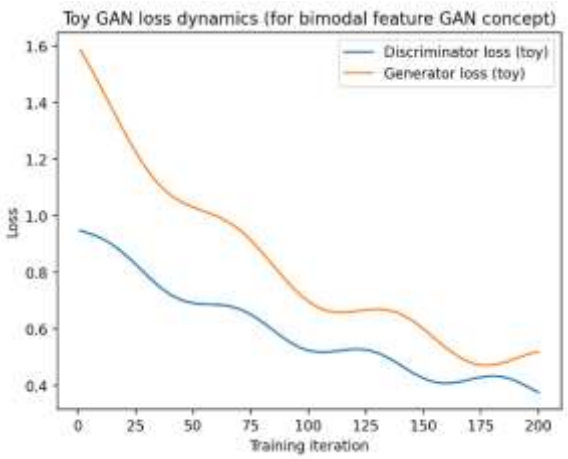
$$\mathcal{L}_G = -\mathbb{E}_{\epsilon\sim p(\epsilon)}\big[\log D\big(G(\epsilon)\big)\big]$$

3.4 Conditional / bimodal GAN for document features

To incorporate inputs (image/text):

$\tilde{z} = G(x_v, x_t, \epsilon)$ $\mathcal{L}_D = -\mathbb{E}_z[\log D(z \mid x_v, x_t)] - \mathbb{E}_\epsilon\big[\log\big(1 - D(G(x_v, x_t, \epsilon) \mid x_v, x_t)\big)\big]$

ISSN: 3067-4190
JOURNAL OF ARTIFICIAL INTELLIGENCE AND BIG DATA DISCIPLINES
Volume: 03 Issue: 03
RECEIVED: JULY 22 REVISED: AUGUST 06 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 04



6. Generative Models for Financial Document Features

User-supplied abstract: Generative and agentic AI enhance a framework for secure feature extraction from financial documents. Partially enabled for edge-cloud processing, the concept addresses concerns about privacy, presence, and data management.
Generative and agentic AI reinforce a framework for secure feature extraction from financial documents. Partially enabled for edge-cloud processing, the mechanism addresses concerns about privacy, presence, and data management. Bimodal Generative Adversarial Networks extract document features from visual and textual channels, encrypting semantically enriched latent representations. Ontology-based representation learning generates data for existing embeddings while securing model training against risk exposure. Financial institutions may apply the technology to fulfil Know Your Customer requirements while mitigating privacy concerns.
Secure processing of financial documents remains a long-standing challenge in edge-cloud computing ecosystems. Generative understanding of financial activities, agentic augmentation through business rules, and privacy-preserving operations partly fulfil the requirements for complete feature extraction. Inadequate security against privacy exposure in traditional systems undermines customer trust. Two mechanisms demonstrate the potential of Generative and Agentic AI. Bimodal Generative Adversarial Networks extract complementary features from the visual and textual channels of financial documents, generating semantically enriched representations while limiting adversarial data disclosure. An ontology for financial activities supplies information for existing embedding sets, facilitating secure model training without direct access to sensitive data. Application for Know Your Customer fulfilment offers a probable market.

6.1. Representation Learning for Documents

Templates-based generation can overcome the data-sparseness problem for many document types, for example, invoices or tickets. For these documents, generative models enable full data-feature extraction from only a few outer-element samples. However, the functional limitation of such document types causes the need for feature representations. A corresponding semistructured-template-based data-generation model extracts outer elements and fills inner elements with semantic-enrichment representations from public resources or large language models based on proposed Task-1 prompts. Directed LSTM-Fully Connected Neural Networks with attention mechanism enable document-style adaptation for multiple semistructured-template-based document types and allow semantic enrichment of the syntactic-grammar representation—also with semantic-inference features. Rich semantic properties of document types have been accumulated in existing ontologies and knowledge, bases. Ontologies create machine-readable semantic descriptions. The existence of many document types enables the use of agentic-principal models to enrich, refine, explain, or falsify document-template layouts and elements such as semistructured, rules, or contents. Supervised representation-learning networks with good generalization capabilities map document representations into aligned spaces to connect summary or description documents with detailed documents. Text, table, or chart are the three main types of elements in documents.

6.2. Ontologies and Semantic Enrichment

ISSN: 3067-4190
JOURNAL OF ARTIFICIAL INTELLIGENCE AND BIG DATA DISCIPLINES
Volume: 03 Issue: 03
RECEIVED: JULY 22 REVISED: AUGUST 06 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 04

Rather than developing a new ontology from scratch, existing ones are reused and linked together to enable semantic enrichment of financial document features with domain-specific information. GermaNet, MindNet, OMCS, WordNet, FrameNet, the DOD ontologies, DBpedia, and OpenCyc have been selected, since they encompass an adequate proportion of lexicon entries related to finance. The integration and linking of these ontologies is achieved through a modular structure in a Web-based format based on OWL-DL. The semantic enrichment process is implemented as a distributed program running on a local computer connected to the Internet.

Semantic enrichment is done using the SA-orc2 algorithm, which matches potentially ambiguous words against the ontologies of the linked set and suggests one or more meanings. Furthermore, SAPhrase_2-Matching is used to enrich named entities with commodity information from DBpedia. The enrichment of other phrases and associations is performed by enriching syntactical constituents of the Focused_3 _Grammatical relations. Enriched features are stored in a hypergraph to facilitate paths through the feature space. These paths are designed to be specific to the task of information extraction of financial documents, enabling answering Who is doing which Action to what or whom with which Object at which Time related to where or why questions.

Method	What is transmitted?	End-to-end latency model
Cloud OCR baseline	Upload full raw doc	$T_{\text{cloud}}(S) = \frac{S(1+o_{\text{enc}})}{B_{\uparrow}} + \frac{S}{r_{\text{ocr}}} + \frac{\gamma S}{B_{\downarrow}}$
Edge generative features	Upload only feature payload	$T_{\text{edge}}(S) = \frac{S}{r_{\text{gen}}} + \frac{\gamma S(1+o_{\text{enc}})}{B_{\uparrow}}$
Hybrid	Upload only ρ fraction for OCR + features back	$T_{\text{hyb}}(S) = \frac{S}{r_{\text{gen}}} + \frac{\rho S(1+o_{\text{enc}})}{B_{\uparrow}} + \frac{\rho S}{r_{\text{ocr}}} + \frac{\gamma S}{B_{\downarrow}}$

7. Conclusion

Generative and agentic AI technologies have the potential to transform business processes in a bottom-up manner by reducing development costs and democratizing access to state-of-the-art technologies. These technologies impact all information-and-product-service-creation sectors and continue to flourish across cloud and edge environments. However, beside programmatically engineered automated agents supported by advanced features such as conversation, common sense reasoning, and multimodality, security and privacy issues also arise. Although generative models of various nature can enhance business processes at multiple levels, the possible introduction of malware, propaganda, and fake news hidden in generated images or video strongly questions the recent capability and model distribution policies of the major generative model creators. To mitigate these security, safety, and privacy issues, a framework for secure feature extraction from financial documents at edge and cloud levels is proposed. The conceptual foundation is based on three aspects. Generative AI and agentic AI support secure document feature extraction. The creation of a stable and secure financial document feature extractor by a programmatically engineered automated agent supports financial services at the edge level without private content exposure. Financial document features are categorical representations of original files that keep public information accessible without disclosing private content. Edge-layer attributes coalesce into the representation of a full document for further processing at the cloud layer, where bids for feature reconstruction are checked against access-control policies

ISSN: 3067-4190
JOURNAL OF ARTIFICIAL INTELLIGENCE AND BIG DATA DISCIPLINES
Volume: 03 Issue: 03
RECEIVED: JULY 22 REVISED: AUGUST 06 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 04

and satisfied by the generalization capability of a generative model.

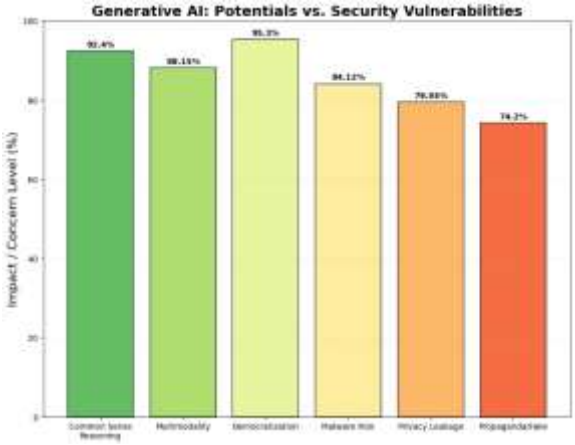


Fig 4: Generative AI: Potentials vs. Security Vulnerabilities

7.1. Final Thoughts and Future Directions

Generative and Agentic AI-Enhanced Feature Extraction for Secure Financial Document Processing in Edge-Cloud Ecosystems

Advancements in generative and agentic artificial intelligence (AI) now enable people to reconsider the long-held belief that only visible data is of value. Recent developments in secured networks, web participatory mechanisms, and self-aware information systems demonstrate that provable security for security-critical environments using cryptographic processes can be achieved without degrading the utility of sensitive information. Building upon these developments, a conceptual framework for secure financial-document feature extraction in edge-cloud ecosystems is proposed. Security is built into the financial-document processing technique in a decentralized manner, leveraging digital transformation characteristics while being applicable across all layers of an edge-cloud architecture. The resulting generative and agentic methods improve feature extraction and enhance security for financial-documents users.

Rapid technological growth, especially in generative-and-agentic artificial intelligence (AI), machine learning, and blockchain, have changed the manner in which society accesses and utilizes information. Generative AI can create meaningful and task-specific synthetic data while agentic AI can autonomously perform complex tasks on behalf of its users. Following the long-held belief that only visible data is of value, people have been reluctant to share the sensitive and privacy-critical information required for the meaningful deployment of generative-and-agentic-AI techniques. However, recent advancements in secured networks, web participatory mechanisms, and a new self-aware-information paradigm have overcome this limitation.

8. References

1. Appalaraju, S., Jasani, B., Kota, B. U., Xie, Y., & Manmatha, R. (2021). DocFormer: End-to-end transformer for document understanding. Proceedings of the IEEE/CVF International Conference on Computer Vision, 993–1003.
2. Arbreha, H. G., Hayajneh, M., & Serhani, M. A. (2022). Federated learning in edge computing: A systematic survey. Sensors, 22(2), 450.
3. Sriram, H. K., Challa, K., Gadi, A. L., & Singreddy, S. (2025). AI and Cloud-Driven Transformation in Finance, Insurance, and the Automotive Ecosystem: A Multi-Sectoral Framework for Credit Risk, Mobility Services, and Consumer Protection. Anil Lokesh and singreddy, Sneha, AI and Cloud-Driven Transformation in Finance, Insurance, and the Automotive Ecosystem: A Multi-Sectoral Framework for Credit Risk, Mobility Services, and Consumer Protection (March 15, 2025).
4. Brecko, A., Kajati, E., Koziorek, J., & Zolotova, I. (2022). Federated learning for edge computing: A survey. Applied Sciences, 12(18), 9124.
5. Chen, Z., Chen, W., Smiley, C., Shah, S., Borova, I., Langdon, D., Moussa, R., Beane, M., Huang, T.-H., Routledge, B., & Wang, W. Y. (2021). FinQA: A dataset of numerical reasoning over financial data. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, 3697–3711.
6. Hong, T., Kim, D., Ji, M., Hwang, W., Nam, D., & Park, S. (2022). BROS: A pre-trained language model

focusing on text and layout for better key information extraction from documents. Proceedings of the AAAI Conference on Artificial Intelligence, 36(10), 10767–10775.

7. Seenu, A., Sheelam, G. K., Motamary, S., Meda, R., Koppolu, H. K. R., & Inala, R. (2025). AI-Driven Innovations in Infrastructure Management with 6G Technology. In 2025 2nd International Conference on Computing and Data Science (ICCDS) (pp. 1–6). IEEE. 2025 2nd International Conference on Computing and Data Science (ICCDS). https://doi.org/10.1109/iccds64403.2025.11209649
8. Huang, Y., Lv, T., Cui, L., Lu, Y., & Wei, F. (2022). LayoutLMv3: Pre-training for document AI with unified text and image masking. Proceedings of the 30th ACM International Conference on Multimedia, 4083–4091.
9. Lee, C.-Y., Li, C.-L., Dozat, T., Perot, V., Su, G., Hua, N., Ainslie, J., Wang, R., Fujii, Y., & Pfister, T. (2022). FormNet: Structural encoding beyond sequential modeling in form document information extraction. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 3735–3754.
10. Narasareddy Annapareddy, V., Pamisetty, A., Malempati, M., Kaulwar, P. K., & Bhardwaj Komaragiri, V. (2025, April). Enhancing Solar Power System Efficiency Through AI-Driven Predictive Maintenance and Cloud-Based Infrastructure Stability Solutions. In International Conference on Smart Computing and Informatics (pp. 328-337). Cham: Springer Nature Switzerland.
11. Lee, K., Joshi, M., Turc, I. R., Hu, H., Liu, F., Eisenschlos, J. M., Khandelwal, U., Shaw, P., Chang, M.-W., & Toutanova, K. (2023). Pix2Struct: Screenshot parsing as pretraining for visual language understanding. Proceedings of the 40th International Conference on Machine Learning, 202, 18893–18912.
12. Li, J., Xu, Y., Cui, L., & Wei, F. (2022). MarkupLM: Pre-training of text and markup language for visually rich document understanding. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 6078–6087.
13. Challa, S. R., Burugulla, J. K. R., Pamisetty, A., Challa, K., & Paleti, S. (2025, April). AI and ML-Powered Cybersecurity Strategies for Cloud Computing: Ensuring Infrastructure Stability in Financial and Retail Sectors. In International Conference on Smart Computing and Informatics (pp. 315-327). Cham: Springer Nature Switzerland.
14. Li, J., Xu, Y., Lv, T., Cui, L., Zhang, C., & Wei, F. (2022). DiT: Self-supervised pre-training for document image transformer. Proceedings of the 30th ACM International Conference on Multimedia, 3530–3539.
15. Li, Q., Wen, Z., Wu, Z., Hu, S., Wang, N., Li, Y., Liu, X., & He, B. (2023). A survey on federated learning systems: Vision, hype and reality for data privacy and protection. IEEE Transactions on Knowledge and Data Engineering, 35(4), 3347–3366.
16. Singireddy, S., Pandiri, L., Motamary, S., Kummari, D. N., Somu, B., & Lakkarasu, P. (2025, August). Blockchain-Powered Claims Validation for Enhancing Trust in Health and Auto Insurance Ecosystems. In International Conference on Artificial Intelligence: Theory and Applications (pp. 26-38). Cham: Springer Nature Switzerland.
17. Li, Y., Qian, Y., Yu, Y., Qin, X., Zhang, C., Liu, Y., Yao, K., Han, J., Liu, J., & Ding, E. (2021). StrucTexT: Structured text understanding with multi-modal transformers. Proceedings of the 30th International Joint Conference on Artificial Intelligence, 1383–1389.
18. Masry, A., & Hajian, A. (2024). LongFin: A multimodal document understanding model for long financial domain documents. Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, 1–14.
19. Mathew, M., Karatzas, D., & Jawahar, C. V. (2021). DocVQA: A dataset for VQA on document images. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2200–2209.
20. Challa, S. R., Kaulwar, P. K., Koppolu, H. K. R., Adusupalli, B., & Suura, S. R. (2025, April). Big Data and AI in Wealth Management: Leveraging Cloud Computing for Secure and Scalable Financial

Infrastructure. In International Conference on Smart Computing and Informatics (pp. 307-318). Cham: Springer Nature Switzerland.

21. Pfitzmann, B., Auer, C., Dolfi, M., Nassar, A. S., & Staar, P. (2022). DocLayNet: A large human-annotated dataset for document-layout segmentation. Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 3743–3751.
22. Tang, Z., Yang, Z., Wang, G., Fang, Y., Liu, Y., Zhu, C., Zeng, M., Zhang, C., & Bansal, M. (2023). Unifying vision, text, and layout for universal document processing. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 19254–19264.
23. Davuluri, P. S. L. (2023). AI-Augmented Sanctions Screening: Enhancing Accuracy and Latency in Real Time Compliance Systems. AI-Augmented Sanctions Screening: Enhancing Accuracy and Latency in Real Time Compliance Systems (December 15, 2023).
24. Xu, Y., Xu, Y., Lv, T., Cui, L., Wei, F., Wang, G., Lu, Y., Florencio, D., Zhang, C., Che, W., Zhang, M., & Zhou, L. (2021). LayoutLMv2: Multi-modal pre-training for visually-rich document understanding. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2579–2591.
25. Zhu, F., Lei, W., Huang, Y., Wang, C., Zhang, S., Lv, J., Feng, F., & Chua, T.-S. (2021). TAT-QA: A question answering benchmark on a hybrid of tabular and textual content in finance. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 3277–3287.
26. Vadisetty, R., Nuka, S. T., Kalisetty, S., Pandugula, C., Burugulla, J. K. R., & Annapareddy, V. N. (2025, February). Generative AI for Advanced Recycling Processes in Polyethylene and Polypropylene Manufacturing. In International Ethical Hacking Conference (pp. 269-284). Singapore: Springer Nature Singapore.
27. Xu, R., Baracaldo, N., & Joshi, J. (2021). Privacy-preserving machine learning: Methods, challenges and directions. arXiv preprint arXiv:2108.04417.
28. Singireddy, J., Kaulwar, P. K., Somu, B., Meda, R., Dodda, A., & Yellanki, S. K. (2025, August). Reinforcement Learning-Based Asset Allocation in Algorithmic Trading for Banking Institutions. In International Conference on Artificial Intelligence: Theory and Applications (pp. 357-369). Cham: Springer Nature Switzerland.
29. Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 37(3), 50–60.
30. Dutta, P., Adusupalli, B., Koppolu, H. K. R., Dodda, A., Yağanoğlu, M., Banerjee, J. S., & Chakraborty, A. (2025, March). Textual Social Data Disinformation Analysis Using a Hybrid Context-Enhanced Deep Learning Model. In Doctoral Symposium on Human Centered Computing (pp. 342-352). Singapore: Springer Nature Singapore.
31. Gill, S. S., Golec, M., Hu, J., Xu, M., Du, J., Wu, H., Walia, G. K., Murugesan, S. S., Ali, B., Kumar, M., Ye, K., Verma, P., Kumar, S., Cuadrado, F., & Uhlig, S. (2024). Edge AI: A taxonomy, systematic review and future directions. Internet of Things and Cyber-Physical Systems, 4, 1–25.
32. Cao, K., Chen, X., & Chen, J. (2023). Edge AI: A survey. Internet of Things and Cyber-Physical Systems, 3, 71–92.
33. Lee, Y., Park, S., & Kim, J. (2023). Deep learning-based document image analysis for intelligent document processing. Expert Systems with Applications, 213, 118958.
34. KOTHAPALLI, S. L., SEENU, A., DILEEP, V., & YASMEEN, Z. (2025). THE FUTURE OF CUSTOMER ENGAGEMENT IN RETAIL BANKING: EXPLORING THE POTENTIAL OF AUGMENTED REALITY AND IMMERSIVE TECHNOLOGIES. INTERNATIONAL JOURNAL, 73(1), 72-79.
35. Zhang, X., Wei, Y., Yang, Q., & Li, J. (2021). Document image understanding with multimodal

ISSN: 3067-4190
JOURNAL OF ARTIFICIAL INTELLIGENCE AND BIG DATA DISCIPLINES
Volume: 03 Issue: 03
RECEIVED: JULY 22 REVISED: AUGUST 06 ACCEPTED: AUGUST 24 PUBLISHED: SEPTEMBER 04

transformer architectures. Pattern Recognition, 119, 108087.
36. Wang, X., Yu, F., Wang, B., & Zhang, Y. (2022). Transformer-based document information extraction: A survey. Information Processing & Management, 59(6), 103092.
37. Ande, R., Mehta, D., Pandugula, C., Krishna AzithTejaGanti, V., & Kalisetty, S. (2025). Modeling the Som Kamla Amba sub-watershed using Artificial Neural Networks (ANN) and the SWAT tool. Available at SSRN 5341755.
38. Li, Y., Huang, Y., Lv, T., Cui, L., & Wei, F. (2022). Layout-aware document understanding with multimodal pre-trained transformers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(12), 9234–9248.
39. Singh, A., Ehtesham, A., & Kumar, S. (2024). Intelligent document processing using deep learning and transformer-based architectures. Journal of King Saud University—Computer and Information Sciences, 36(2), 101–114.
40. Maguluri, K. K. (2025). Ethical challenges in artificial intelligence. How Artificial Intelligence is Transforming Healthcare IT: Applications in Diagnostics, Treatment Planning, and Patient Monitoring, 132.
41. Toprak, A., & Turan, M. (2024). Transformer-based approach for automatic semantic financial document verification. IEEE Access, 12, 184327–184349.
42. Serbanescu, V.-A., & Dhali, M. (2025). Deep learning for effective classification and information extraction of financial documents. Proceedings of the 2025 International Conference on Computer Science and Information Technology, 1–8.
43. Wang, W., Hu, H., Zhang, Z., Li, Z., Shao, H., & Dahlmeier, D. (2025). Document intelligence in the era of large language models: A survey. arXiv preprint arXiv:2510.13366.
44. Pandugula, C., Ganti, V. K. A. T., Kalisetty, S., Ande, R., & Mehta, D. (2025). Modeling the Som Kamla Amba Sub-Watershed Using Artificial Neural Networks (ANN) and the SWAT Tool. Available at SSRN 5341614.
45. Abdi, M. J., & Baghshah, M. S. (2023). Efficient document understanding with vision-language transformers. Neural Processing Letters, 55, 10457–10478.
46. Li, C., Li, R., Zhang, S., & Wang, X. (2023). Multimodal transformer networks for financial document understanding. IEEE Transactions on Artificial Intelligence, 4(6), 1548–1561.
47. Zhang, Y., Liu, Y., & Jin, L. (2021). Document layout analysis: A comprehensive survey. ACM Computing Surveys, 54(8), 1–36.
48. Xu, Y., Lv, T., Cui, L., Wang, G., Lu, Y., Florencio, D., Zhang, C., & Wei, F. (2022). XFUND: A benchmark dataset for multilingual visually rich form understanding. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 207–218.
49. ADUSUPALLI, B. (2025). Redefining Financial Risk Strategies: The Integration of Smart Automation, Secure Access Systems, and Predictive Intelligence in Insurance, Lending, and Asset Management. JAIBDD.
50. Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: A pre-trained biomedical language representation model for biomedical text mining. Bioinformatics, 36(4), 1234–1240.
51. Howard, J., & Ruder, S. (2020). Universal language model fine-tuning for text classification. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 1–12.
52. Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R. G. L., Eichner, H., El Rouayheb, S., Evans, D., Gardner, M., Garrett, Z., Gascón, A., Ghazi, B., Gibbons, P. B., … Zhao, S. (2021). Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14(1–2), 1–210.
53. Murshed, M. G., Murphy, C., Hou, D., Khan, N., Ananthanarayanan, G., & Hussain, F. (2020). Machine learning at the network edge: A survey. ACM Computing Surveys, 53(5), 1–37.

54. Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2020). Edge intelligence: Paving the last mile of artificial intelligence with edge computing. Proceedings of the IEEE, 107(8), 1738–1762.
55. Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2020). Deep learning with differential privacy: Privacy-preserving model training and inference. ACM Transactions on Privacy and Security, 23(4), 1–36.
56. Zhu, L., Liu, Z., & Han, S. (2020). Deep leakage from gradients. Advances in Neural Information Processing Systems, 33, 14774–14784.
57. Geiping, J., Bauermeister, H., Dröge, H., & Moeller, M. (2020). Inverting gradients—How easy is it to break privacy in federated learning? Advances in Neural Information Processing Systems, 33, 16937–16947.
58. Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., Kiddon, C., Konečný, J., Mazzocchi, S., McMahan, H. B., Van Overveldt, T., & Roselander, J. (2020). Towards federated learning at scale: System design. Proceedings of Machine Learning and Systems, 2, 374–388.
59. Zhou, Y., Yu, Y., Wang, Z., & Li, H. (2024). Efficient and secure edge intelligence: A survey of model compression, privacy, and deployment. IEEE Internet of Things Journal, 11(8), 13452–13471.
60. Vishnubhatla, S. (2025). Hybrid intelligence for information management systems: Converging edge AI and cloud for real-time document understanding. SSRN Electronic Journal.