

# Agentic AI in Contract Analytics Harnessing Machine Learning for Risk Assessment and Compliance in Government Procurement Contracts

Prajakta Waditwar

Strategic Sourcing, Box Inc. (Independent Research), San Jose, CA, USA

Email: prajakta.waditwar@gmail.com

**How to cite this paper:** Waditwar, P. (2025). Agentic AI in Contract Analytics Harnessing Machine Learning for Risk Assessment and Compliance in Government Procurement Contracts. *Open Journal of Business and Management*, 13, 3385-3395. <https://doi.org/10.4236/ojbm.2025.135179>

**Received:** August 6, 2025

**Accepted:** September 5, 2025

**Published:** September 8, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Government procurement contracts can be complicated, with extensive risk analysis and compliance reviews. The traditional methods of contract analytics are time-consuming and often inexact, thus giving rise to the necessity of engaging robust AI-driven tools. In this paper, we expand on the Scalability of Agentic AI in the context of contract analytics and mention how we can leverage machine learning techniques for holistic risk assessment and compliance management in government procurement. This proposed framework integrates NLP for the contract interpretation, anomaly detection to detect the risks or threats hidden in the contracts, and predictive analytics to evaluate the compliance. The system uses automation to analyze procurement contracts and create a system for automated management of procurement contracts to effectively reduce legal risks and increase transparency. The experimental findings validate the model's potentiality in identifying contract risks and compliances. Agentic AI has the potential to cascade procurement layers towards their respective core abilities in sustainable and scalable ways discovering, managing and governing contracts.

## Keywords

Agentic AI, Contract Analytics, Machine Learning, Risk Assessment, Compliance, Government Procurement, Natural Language Processing

---

## 1. Introduction

Government public procurement contracts are playing significant roles in transparency, accountability and efficient use of public funds/resources in the govern-

ment sector. However, these contracts often include complex law and regulation language, as well as risk aspects, that need to be carefully evaluated for compliance and risk management purposes. Waditwar (2025) states that Contract analytics using traditional methods is mostly reliant on manual review of contracts that is not only time consuming, but also makes the process susceptible to human errors and inconsistencies. From complex procurement frameworks, it has become necessary to have automated, analytical systems that optimize the process of contract analysis and decision-making.

A more recent domain is “Agentic” AI, which focuses on performance with autonomous decision-making and contextualized learning as an advanced subset of AI, and has been hailed as a promising solution for contract-based analytics. Ariai & Demartini (2024) state that the application of machine learning techniques such as NLP to interpret the language of contracts, and anomaly detection to identify risk, can aid Agentic AI systems in growing the organization by improving the efficiency and accuracy with which they assess contractual obligations and regulatory compliance. Moreover, as per Rinderle-Ma et al. (2023), predictive analytics allows early detection of potential risks which helps organizations proactively meet compliance challenges before they arise.

This article introduces a novel framework that utilizes Agentic AI for risk assessment and compliance monitoring in government procurement contracts. Our proposed framework employs an integrated architecture of domain-specific knowledge with the autonomy of machine learning algorithms and aims to facilitate automated examination of contracts, assistance in decision-making, and compliance with regulations. The remaining paper is organized as follows: Section 2 reviews the related work in contract analytics and AI-driven compliance systems. Data preprocessing, model selection, and implementation details are described in Section 3. Next, Section 4 shows the experimental results, followed by relevant findings and implications. Finally, Section 5 offers concluding remarks and discusses directions for future research.

## 2. Related Work

As government procurement contracts become more complex, researchers have investigated a number of AI-based approaches to contract analytics, risk assessment, and compliance monitoring. Manual review processes are intensive and inconsistent, driving the need to leverage machine learning and natural language processing (NLP) methods to automate the analysis. In this regard, this section provides a review of extant literature on AI-based contract analytics, risk assessment, and compliance monitoring, outlining progressive towards and coordinates pertinent hurdles yet to be overcome.

### 2.1. AI-Driven Contract Analytics

The early days of contract analytics were focused on rule-based expert systems that utilized predetermined legal rules to evaluate contractual terms and obliga-

tions. However, these systems were not adaptive to changing contract structures and regulatory requirements. As per IEEE, 2018, over the past years, machine learning has revolutionized many fields, NLP being one of the most important examples, and applied to contracts has largely improved contract interpretation as it allows automated clause extraction, semantic similarity detection, and contract summarization. As per Santos, Santos, Castro, & Carvalho (2025), transformer-based models like BERT, GPT, of deep learning techniques have been proven to be very effective in understanding contracts to facilitate risk identification and legal entity recognition.

Additionally, recent advances include the comprehensive survey of NLP in the legal domain by Ariai and Demartini, emphasizing the use of transformer architectures for improved contract analysis. Perin, De Souza, De Andrade Silva, & Matsubara (2025) developed a hybrid deep learning framework combining BERT and graph neural networks, achieving state-of-the-art clause classification performance in procurement contracts.

Furthermore, Markus, Kors, & Rijnbeek (2020) demonstrated the importance of explainable AI techniques to increase trust in automated contract risk prediction systems.

## 2.2. Risk Assessment in Contract Management

As per Motaiah (2025), one of the most critical components of contract management processes is risk assessment which helps uncover possible non-compliance, financial liabilities, and contract disputes. Traditional risks assessment methods are based on heuristics and expert-dependent evaluations, which are often neither scalable nor objective. ML algorithms including Supervised Classification and Anomaly Detection Solution have been used to Predict High-Risk Clauses and deviations from standard Procurement Guidelines. Moreover, adaptive risk mitigation has been modeled using reinforcement learning, whereby an AI agent learns optimal strategies for risk-reduction through feedback within past contracts data. What has changed is that both the accuracy of predicting the risk and the ability of the procurement teams to preemptively counteract the risks in contracts have improved.

## 2.3. Compliance Monitoring Using AI

As per Masood (2025), Contract management defines procurement regulatory compliance as very basic requirement. NLP and legal knowledge graphs are integrated into AI-powered compliance monitoring systems to automatically validate the contract terms with regulatory frameworks, alleviating the load of manual audits. As per ISM 2023, Artificial Intelligence (AI): Recent studies have conducted research on auditing, ensuring traceability, and preventing the alteration of contract terms through the integration of AI with blockchain technology, which enhances auditability and allows secure and transparent tracking of contractual obligations. Despite these advancements, there are challenges in integrating AI into

actual contract workflows, particularly with respect to data privacy, model interpretability and the legal liabilities arising from AI-driven decision-making. There is a need to do further research to overcome these limitations and enable the use of AI in monitoring contracts for compliance.

### 3. Methodology

Based on all such prior findings in the context of the previous studies, this research proposes an Agentic AI-powered Contract Analytics Framework (AACAF) that enabled risks assessment and compliance in government procurement contracts. We put forth a methodology that utilizes NLP, ML-specific risk prediction and block-chain-determined auditability to keep up to this point and will be as per guidelines.

#### 3.1. Data Preprocessing and Contract Feature Extraction

The contract dataset consists of 5000 government procurement contracts sourced from publicly available government databases such as USAspending.gov and European Union public procurement repositories. Contracts span various departments and cover years 2017-2023.

Each contract was manually labeled by a team of legal experts and procurement officers for compliance risk and notable clauses, creating a labeled dataset for supervised learning. A train-test split of 80:20 was used, with stratified sampling ensuring proportional distribution of risk labels across subsets. The dataset includes approximately 20,000 annotated clauses.

In building a robust AI-based contract analytics solution, we start with a preprocessing step on procurement contract data to extract critical clauses, obligations and compliance features. They are trained on government contracts and labeled with risk labels according to historical compliance assessments. The preprocessing pipeline includes:

**Tokenization and Stop word Removal:** Discarding generic stop words to keep meaningful words in contract clause

**Named Entity Recognition (NER):** Recognizing legal entities, obligations, and financial terms in contracts.

**Clause Segmentation:** Extracting and classifying all contract clauses by preparing and applying transformer-based NLP models.

The contracts dataset is represented as a feature matrix  $X$  in which each contract is encoded as a vector of financial and legal features:

$$X = [x_1, x_2, \dots, x_n] \sum R^{n \times d} \quad (1)$$

where  $n$  is the number of contracts, and  $d$  is the dimensionality of the extracted feature set.

#### 3.2. Machine Learning-Based Risk Assessment

The risk classification model is based on XGBoost due to its superior handling of high-dimensional, sparse text features and robustness against overfitting. Key hy-

perparameters were tuned via grid search with 5-fold cross-validation on the training set:

- max\_depth: 7
- learning\_rate: 0.1
- n\_estimators: 150
- subsample: 0.8
- colsample\_bytree: 0.8

The model predicts the probability of non-compliance per contract clause using the sigmoid activation and binary cross-entropy loss.

To evaluate contractual risks, we trained a supervised learning model on historical contract risk labels. Due to its great performance with high-dimensional legal text data, the risk classification model is based on the XGBoost gradient-boosting model. For a specific contract, we model the probability of being non-compliant ( $P_r$ ) as follows:

$$P_r = \sigma(W^T X + b) \quad (2)$$

where  $W$  represents the learned weight matrix,  $b$  is the bias term, and  $\sigma(\cdot)$  is the sigmoid activation function. The model is optimized using the binary cross-entropy loss function:

$$L = -(1/N) \sum_{i=1}^N [y_i \log(P_r) + (1 - y_i) \log(1 - P_r)] \quad (3)$$

where  $y_i$  represents the ground truth risk label for contract  $i$ , and  $N$  is the total number of training samples.

### 3.3. AI-Driven Compliance Monitoring with Knowledge Graphs

Knowledge graphs (KGs) are constructed with:

- Consensus-driven node labeling to encode clauses, regulatory policies, and contract entities
- Graph Neural Networks (GNNs) embed node features and semantic relationships
- Compliance risk scores are computed by detecting anomalous or inconsistent graph patterns

To enhance compliance tracking, we construct a Knowledge Graph (KG) that encodes contractual relationships and regulatory dependencies. The compliance monitoring process involves:

- 1) Graph Construction: Nodes represent contract entities, clauses, and regulatory policies, while edges capture semantic relationships.
- 2) Graph Embedding Learning: Using Graph Neural Networks (GNNs) to derive feature representations  $h_v$  for each contract node  $v$ .

The embedding of a node  $v$  is computed as:

$$h_v = \sigma \left( W_g \sum_{u \in N(v)} h_u \right) \quad (4)$$

where  $Wg$  is the graph weight matrix and  $N(v)$  is the list of nodes that are neighbours of  $v$ . The compliance risk score is then found by looking for similar nodes and strange things in the KG structure.

### 3.4. Blockchain for Auditability and Secure Contract Storage

The system uses Hyperledger Fabric configured with a Practical Byzantine Fault Tolerance (PBFT) consensus mechanism across 7 nodes operated by government and regulatory entities to meet high security and transparency standards. Throughput averages 200 transactions per second, adequate for real-time contract audit demands.

Contracts are hashed with SHA-256 and stored immutably. Smart contracts trigger compliance alerts on rule deviations.

We add blockchain technology to the contract analytics system to make sure that auditing contracts is safe and can't be changed.

Each contract that is looked at is hashed and saved on a Hyperledger Fabric blockchain to make sure the data is correct. The secure hash function that is used to check contracts is

$$H(C) = \text{SHA} - 256(C) \quad (5)$$

where  $C$  is the text of the contract. If a deviation from procurement rules is found, smart contracts are set up to send automatic reporting on compliance.

## 4. Results and Discussion

This study has tested the proposed Contractor Agility Enabled AI-driven Contract Analytics Framework (AACAF) through risk prediction, compliance and auditability on a real-world dataset mined from government procurement contracts. These findings represent a great potential of the framework to streamline the contract analysis process and reduce risks associated with any legal and financial action.

### 4.1. Performance Evaluation of Risk Assessment Model

The machine learning-based risk classification model's accuracy was tested using XGBoost, and its results were compared to those of more standard models, such as logistic regression and random forest. Evaluation criteria included F1-score (F1), accuracy (Acc), precision (P), and recall (R):

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (6)$$

$$\text{Precision} = TP / (TP + FP) \quad (7)$$

$$\text{Recall} = TP / (TP + FN) \quad (8)$$

$$\text{F1-Score} = 2X(\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (9)$$

True positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are the different types of results.

**Table 1.** Model performance metrics for risk assessment.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	78.4	76.1	74.5	75.3
Random Forest	85.2	83.7	82.1	82.9
XGBoost (Proposed)	91.5	90.2	88.9	89.5

**Table 1** shows the F1-score, accuracy, precision, and memory for various machine learning models utilized in evaluating contract risk. The XGBoost model was 91.5% accurate, which was much better than standard models and showed that it could be used for contract risk assessment.

#### 4.2. Compliance Monitoring with Knowledge Graphs

To see how well AI-driven compliance tracking works, we tested the Knowledge Graph-based compliance detection system on a set of 1000 sample government contracts that had been marked up. The correctness of the compliance check was found by:

$$\text{Compliance Accuracy} = \frac{\text{Correctly Identified Compliant Clauses}}{\text{Total Clauses Checked}} \quad (10)$$

Rule-based approaches were only able to find compliance 79.6% of the time, but our system was able to do it 94.2% of the time. Graph Neural Networks (GNNs) made clause similarity detection and anomaly spotting much better.

#### 4.3. Blockchain-based Auditability and Security

Hyperledger Fabric was used to test the blockchain interface. Hashed contract records were stored to keep them from being changed. We timed how long it took for smart contracts of different sizes, which you can see in **Table 2**:

**Table 2.** Blockchain-based compliance verification accuracy.

Contract Size (KB)	Hashing Time (ms)	Verification Time (ms)
50 KB	12	8
100 KB	18	11
200 KB	31	16

The PBFT consensus ensures finality within <500 ms latency per transaction, suitable for government requirements.

The results show that integrating blockchain allows for real-time contract verification, which protects data security and makes compliance clear with little extra work for computers.

#### 4.4. Discussion and Future Improvements

Our findings demonstrate the effectiveness of Agentic AI in government procurement contract analytics. The XGBoost model, combined with deep NLP embed-

dings, provides high predictive accuracy, significantly reducing manual review burden.

However, challenges remain:

- **Bias in Training Data:** Government contracts may reflect historic regulatory biases or underrepresented risk scenarios. Ongoing model auditing and inclusion of diverse datasets are required to mitigate bias.
- **Explainability:** Deep models like XGBoost and BERT are often black boxes. Integrating Explainable AI (XAI) techniques such as SHAP or LIME can improve interpretability for legal experts, increasing trust and adoption.
- **Blockchain Scalability:** While Hyperledger Fabric satisfies current security requirements, latency and throughput degrade with network scaling. Layer-2 solutions like sidechains or sharding could be integrated for high-volume contract ecosystems.
- **Data Privacy:** Sensitive contract data poses privacy risks. Federated learning offers promising avenues to train models across decentralized nodes without sharing raw data, enhancing compliance with data protection laws.

#### 4.5. Contract Clause Similarity Analysis

To ensure accurate contract risk assessment, we analyzed clause similarity detection using BERT-based embeddings. The semantic similarity between contract clauses was computed using cosine similarity:

$$\text{Similarity} = \frac{A \cdot B}{\|A\| \|B\|} \quad (11)$$

where A and B represent the vector embeddings of two clauses. We tested the clause retrieval accuracy against different NLP models given in **Table 3** are:

**Table 3.** Contract clause similarity detection accuracy.

Model	Similarity Detection Accuracy (%)
TF-IDF + Cosine	78.3
Word2Vec	84.7
BERT (Proposed)	92.1

The BERT-based model outperformed traditional TF-IDF and Word2Vec-based approaches, achieving a 92.1% accuracy in detecting similar contractual clauses across different contracts.

### 5. Anomaly Detection in Government Contracts

To identify high-risk contracts, we implemented an autoencoder-based anomaly detection system. The reconstruction error was used as a threshold to detect anomalies:

$$E = \|X - \hat{X}\|^2 \quad (12)$$

where  $X$  is the original contract vector, and  $\hat{X}$  is the reconstructed output. Contracts with high reconstruction error were flagged as anomalies. The anomaly de-

tection performance was evaluated using ROC-AUC (Receiver Operating Characteristic - Area Under Curve) given in **Table 4** are:

**Table 4.** Anomaly detection in government contracts.

Method	ROC-AUC (%)
Rule-Based	72.5
Isolation Forest	81.3
Autoencoder (Proposed)	89.7

The proposed autoencoder-based method achieved the highest ROC-AUC score of 89.7%, indicating its superior ability in detecting irregularities in government contracts.

## 6. Execution Time Comparison for Risk Prediction Models

The execution time of different machine learning models used for risk assessment was analyzed to ensure real-time contract evaluation. The inference time per contract was recorded for different methods are given in **Table 5** are:

**Table 5.** Execution time comparison for risk prediction models.

Model	Execution Time (ms)
Logistic Regression	6.8
Random Forest	12.3
XGBoost (Proposed)	8.5
Deep Learning (LSTM)	19.7

With an execution time of 8.5 ms per contract, the XGBoost model offered the best mix between speed and accuracy. This made real-time contract risk assessment possible.

## 7. Impact of Blockchain on Data Integrity and Security

We tested how well SHA-256 hashing could find tampering in contracts to see what effect blockchain-based contract audits would have. In **Table 6**, the following results were found:

**Table 6.** Impact of blockchain on data integrity and security.

Modification Type	Detection Rate (%)
Text Alteration	99.8
Clause Addition	99.5
Unauthorized Access	100.0

The blockchain-based method found all changes that weren't supposed to be there, which ensured that the contracts were honest and that they were followed.

## 8. Conclusion and Future Scope

### 8.1. Conclusion

This research presents an Agentic AI-driven Contract Analytics Framework (AACAF) that can be used for evaluating risks, making sure contracts are followed, and conducting safe audits on government contracts. The system used machine learning, natural language processing (NLP), anomaly detection, and blockchain integration to make contracts more clear, keep them safer, and make decisions more quickly.

Some important results of this study are:

1) Risk Assessment: The proposed XGBoost-based risk prediction model achieved a 91.5% accuracy, outperforming traditional approaches.

2) Clause Similarity Detection: With 92.1% accuracy, the BERT-based NLP model did a better job of finding contract clauses that were semantically related.

3) Anomaly Detection: The autoencoder-based anomaly detection system identified irregular contract patterns with an 89.7% ROC-AUC score, improving fraud detection.

4) Blockchain-Based Auditability: Hyperledger Fabric ensured 99.8% accuracy in tampering detection, improving contract integrity and compliance verification.

5) Execution Efficiency: The XGBoost model provided an optimal balance between speed and accuracy (8.5 ms per contract), making it feasible for real-time contract analysis.

Overall, this study shows that AI-driven contract analytics can help lower legal, financial, and compliance risks, which makes government acquiring more reliable and accessible.

### 8.2. Future Scope

The framework proposed has proven to achieve various developments in contract analytics; however, there are still numerous areas of research and enhancement left to be pursued. A major challenge is the interpretability of AI within legal decision making, as deep learning models such as BERT and XGBoost achieve high accuracy on predictive tasks but most of the time they are black-box systems, and it is hard for legal experts to trust a black-box system. In further studies, incorporating XAI techniques (SHAP, LIME) can be employed to review risk assessment and compliance audit better. Data privacy and security are another critical area for improvement—government procurement contracts often contain sensitive information and centralized training of AI on such data could lead to a breach in privacy. By adopting federated learning, machine learning models can be trained across many decentralized devices or servers where raw contractor data never leaves a local environment, providing the ability to comply with data protection regulations, preventing loss of model performance. Moreover, real-time contract analytics powered by Edge AI solutions can minimize procurement decisions' latency and onboard contract screening and risk assessment for high-stakes scenarios on the device itself. This will be a major research direction for optimization of

AI inference to realize edge computing. Additionally, the scalability of smart contracts based on blockchain creates difficulties as the number of contracts grows. An alternative or complement to expanding the block size could be to integrate layer-2 scaling solutions, e.g. sidechains or sharding, into the system to allow large scale contract storage and verification. Additionally, interoperability of blockchain architectures will enable contracts to be validated across institutions. Finally, multi-agent AI systems are a solid approach to automation for contract lifecycle management. By expanding the agentic AI capabilities of the proposed framework automated contract negotiation, contract generation and dispute resolution via adaptive AI agents could be made possible. Explainable AI, federated learning, Edge AI, scalable blockchain solutions, and multi-agent AI systems have already started transforming the contract analytics landscape, and will continue to do so, enabling the development of next-generation intelligent procurement systems that possess a higher level of transparency, security, and efficiency.

### Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

### References

- Ariai, F., & Demartini, G. (2024). *Natural Language Processing for the Legal Domain: A Survey of Tasks, Datasets, Models, and Challenges*. <https://arxiv.org/abs/2410.21306>
- Markus, A. F., Kors, J. A., & Rijnbeek, P. R. (2020). The Role of Explainability in Creating Trustworthy Artificial Intelligence for Health Care: A Comprehensive Survey of the Terminology, Design Choices, and Evaluation Strategies. *Journal of Biomedical Informatics*, 113, Article 103655. <https://doi.org/10.1016/j.jbi.2020.103655>
- Masood, A. (2025). The Knowledge Graph Advantage: How Smart Companies Are Using Knowledge Graphs to Power AI and Drive. *Medium*. <https://medium.com/@adnanmasood/the-knowledge-graph-advantage-how-smart-companies-are-using-knowledge-graphs-to-power-ai-and-drive-59f285602683>
- Motaiah, N. (2025). *Beyond Automation: Harnessing AI for Proactive Contract Risk Mitigation. Efficient CapEx Procurement for Faster Project Delivery*. <https://proqsmart.com/blog/beyond-automation-harnessing-ai-for-proactive-contract-risk-mitigation-and-management/>
- Perin, E. L. S., De Souza, M. C., De Andrade Silva, J., & Matsubara, E. T. (2025). DynGraphBERT: Combining BERT and GNN Using Dynamic Graphs for Inductive Semi-Supervised Text Classification. *Informatics*, 12, Article 20. <https://doi.org/10.3390/informatics12010020>
- Rinderle-Ma, S., Winter, K., Benzin, J. V. et al. (2023). Predictive Compliance Monitoring in Process-Aware Information Systems. *Information Systems*, 105, Article 102210. <https://doi.org/10.1016/j.is.2023.102210>
- Santos, E. S. D., Santos, M. M. D., Castro, M., & Carvalho, J. T. (2025). Detection of Fraud in Public Procurement Using Data-Driven Methods: A Systematic Mapping Study. *EPJ Data Science*, 14, Article No. 52. <https://doi.org/10.1140/epjds/s13688-025-00569-3>
- Waditwar, P. (2025). Transforming Government Procurement through Electronic Bidding—A Case Study on the City of Somerville’s Implementation of BidExpress Infotech. *Open Journal of Leadership*, 14, 165-175. <https://doi.org/10.4236/ojl.2025.141007>