International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 12-No.1 (2026) pp. 1273-1279
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Architecting Intelligent Tax Automation: Research Innovations in Machine Learning for Global Compliance

Vedashree Kedar Karandikar*

Independent Researcher, USA
* **Corresponding Author Email:** reachvedashree@gmail.com - **ORCID:** 0000-0002-5247-4472

Article Info:

DOI: 10.22399/ijcesen.5001
Received : 19 December 2025
Revised : 15 February 2026
Accepted : 20 February 2026

Keywords

Transformer-Based Tax Classification,
Item-to-Tax Semantic Classification,
Confidence-Calibrated Automation,
Human-in-the-Loop Governance,
Production-Scale Compliance Systems

Abstract:

The global indirect tax compliance of large-scale digital commerce platforms has become a complex, high-stakes systems problem due to jurisdictional fragmentation, regular change of regulations, and the high pace of expansion of heterogeneous product catalogs. Rule-based tax engines, although auditable and deterministic, fail to scale in such a situation because their authoring processes are fragile, maintenance is expensive, and their semantic knowledge of product data is limited. This article provides a detailed design of an intelligent tax automation system that is based on machine learning-based item-to-tax prediction services, supported by confidence-aware orchestration, human-in-the-loop protection, and explainability features that are appropriate in regulated financial settings. The suggested framework uses transformer-based language models that are trained on large-scale and multilingual commerce data to predict tax classifications directly based on item titles, descriptions, and structured taxonomy cues. Instead of using fixed mappings, the system is trained on semantic associations between product representations and jurisdiction-specific tax treatments, allowing it to correctly process long-tail, ambiguous, and newly added items. Calibrated confidence scores are provided with predictions, and this indicates whether the transactions can be safely automated, sent to policy validation, or sent to expert scrutiny. This is the selective automation model that balances operational efficiency and regulatory risk, compliance integrity, and scale. The architecture is deployed in a controlled machine learning system and combines continuous monitoring, auditability, and retraining pipelines based on feedback. The experience of large-scale deployments has shown that they can substantially reduce the effort required for manual rule formulation and scrutiny, increase the accuracy of classification into thousands of categories, and have a quantifiable financial effect, without compromising transparency to auditors and other regulatory stakeholders. The article defines intelligent, ML-driven tax automation as a feasible and responsible alternative to the legacy rule-based systems in the global compliance areas.

1. Introduction

E-commerce platforms that conduct business across national boundaries are becoming more complicated when it comes to indirect tax compliance. The regimes of value-added tax (VAT), goods and services tax (GST), and sales tax differ in a wide range of jurisdictions with regard to tax rates, product definitions, exemptions, thresholds, and effective dates. The pace of regulatory changes is high and retroactive, which puts a long-term strain on compliance systems to be accurate, explainable, and scalable to audit.Traditionally, the systems of determining taxes have been based on manually written rules and fixed mappings of product types and taxation codes. These systems offer clear control and traceability, which is necessary in regulatory audits. They do not, however, fit well in the contemporary marketplaces where product catalogs are changing very fast, there are inconsistencies in the descriptions of items made by sellers, and there is the emergence of digital and hybrid goods. Every new type of product or change of jurisdiction is manually handled, which creates operational bottlenecks, slows compliance updates, and increases the risk of misclassification.This article argues that intelligent tax automation requires a shift to be made between rule-based systems and learning based systems capable of generalizing between product semantics without reducing regulatory protection. Recent advances in natural language processing, particularly transformer models, enable systems to learn a tax-relevant meaning directly using item data on a marketplace scale. The main issue is not only the accuracy of prediction but also the design of end-to-end systems that would incorporate machine learning into compliance processes with the right governance, transparency, and human control.

2. Problem Characteristics of Global Item-to-Tax Classification

The classification of items to tax is a distinctive machine learning problem because of the character of commerce information and the regulatory environment in which the decision is put into action. Product information is, by its nature, noisy, incomplete, and heterogeneous. The names of items tend to be brief, unorganized, and erratic, and attributes can be absent or filled in incorrectly. Nonetheless, taxability is based on subtle interpretations of product functionality, composition, mode of delivery, and intended use, which differ across jurisdictions.

Online goods and services also make it more difficult to classify. Depending on the jurisdiction-specific definitions and delivery mechanisms, software, subscriptions, digital media, and virtual goods may be taxed as tangible goods, services, or intellectual property. The traditional taxonomies have a low expressiveness to represent these differences, leading to fragile sets of rules and unreliable results at scale.

Tax classification is a risky field as far as systems are concerned. Mistakes may result in regulatory fines, audit results, loss of customer confidence,e and misstatements. Consequently, complete autonomy in decision-making with no transparency and control is not acceptable. Any smart automation system should hence be able to integrate predictive precision with clear governance systems that facilitate auditability and regulatory scrutiny.

3. Tax Classification Architecture based on Machine Learning

The smart tax automation architecture outlined in this paper is deliberately transformer-centric, as per the semantic complexity and scale of digital commerce today. It consists of a prediction service based on transformer-based language models that are trained on large amounts of commerce-specific, multilingual text. These models learn tax-relevant meaning directly based on short, noisy item titles, descriptions, and structured listing attributes, and can generalize well across long-tail and never-seen items.

The system would analyze various heterogeneous inputs to identify the most suitable tax category of a given item. These inputs usually consist of seller-reported product types, unstructured listing text, structured item attributes, and auxiliary signals like images or seller profile indicators, where present. Seller-provided categories are considered to be informative and fallible signals, whereas model predictions are an independent semantic evaluation. An agreement and disagreement are assessed between signals by a configurable decision layer, and category-specific confidence thresholds are used to decide the ultimate classification result. This design is a direct recognition of the real-life situation where sellers might misclassify goods without any intention, models might be uncertain, or both can be wrong. The ultimate conclusion is not a one-model product, but the product of controlled concert.

Previously used methods of tax classification tended to use ensemble machine learning methods- like gradient boosting, random forests, or support vector machines- frequently alongside fixed rule engines. Although useful in structured and stable catalogs, these approaches fail in semantic ambiguity, sparse text, multilingual listings, and long-tail distributions typical of contemporary marketplaces. Transformer-based models overcome these drawbacks by encoding contextual meaning and allowing confidence-aware decision-making at scale.

3.1 Novel Contributions

The contributions that this work makes to the field of compliance automation are as follows:

1. Transformer-Centric Tax Classification Architecture: The introduction of a production-scale, transformer-based item-to-tax classification system that supplants fragile rule authoring with semantic inference.
2. Discrepancy-Aware Decision Governance: Formalization of seller-versus-model disagreement as a first-class systems problem, solved by confidence-aware orchestration as opposed to fixed precedence rules.
3. Selective Automation with Human-in-the-Loop Protection (Calibrated Confidence Routing Design): Design of calibrated confidence routing that offers maximum

automation without compromising regulatory control and auditing.

4. Ongoing Learning in a Controlled Space: Feedback of Expert Review, Audits, and Policy Revisions into Managed Retraining Pipes without compromising Compliance Assurances.

The intelligent tax automation system mentioned in the article will be implemented on the listing and transaction level of large digital marketplaces. Its design consists of, at its core, an item classification engine that relies on machine learning and which considers inputs of a broad range of heterogeneous items to decide what sort of tax rate should be applied to a particular item. Frequently, these inputs are seller-reported product types, unstructured listing text (titles and descriptions), structured item attributes, and supplementary signals like images or seller profile indication, where it exists.

Language models based on transformers constitute the main prediction layer and allow semantic interpretation of short and noisy commerce text. These models are trained with a large amount of historical marketplace data and optimized with expert-validated tax classification results. The system can generalize beyond strict keyword matching and deal with long-tail, ambiguous, and unseen items by acquiring contextual relationships between item descriptions and tax-relevant categories.

The architecture is specifically designed to support multi-source inference and not single signal support. The categories provided by sellers are considered as informative but fallible inputs, and model predictions are used to provide an independent semantic evaluation. An adaptable decision layer compares congruence or incongruence between these signals and uses tuned confidence limits to decide the ultimate classification result. The design is a mirror of real-life situations when sellers can misclassify items accidentally, models can be uncertain, or both sources can be wrong.

3.2 Essence of the transformer-based architecture

The core of the proposed architecture is a prediction service, as the language models, based on transformers, are used to identify the tax-relevant categories of items. These models are trained with a lot of commerce-specific and multilingual text, which enables them to learn semantic patterns that are marketplace-specific. The model is trained on the association between the representations of items and jurisdiction-specific tax treatments by fine-tuning tax outcomes that have proven correct in the past.

Unlike the traditional classifiers, transformer-based models encode contextual meaning, meaning that they can handle synonymy, polysemy, and sparse descriptions because they use fixed keywords or manually designed features. Text embeddings are accompanied by categorical hierarchies, price ranges, features of sellers, and previous behavior to make the predictions more robust. The system is implemented as a service-oriented and modular element of a bigger compliance system. This makes it easy to train and version models independently, deploy, and monitor, which makes it easy to continue to improve without disrupting downstream tax calculation and reporting systems.

4. Confidence-Aware Automation and Human-in-the-Loop Design

In order to achieve responsible automation, the model prediction is accompanied by a calibrated confidence score that reflects the confidence of the system. These scores are applied to define the processing route of each transaction. High-confidence predictions are automatically used, which allows straight-through processing on a large scale. Cases with medium-confidence are passed through policy validation checks, and low-confidence or high-risk cases are sent to expert human reviewers.

This is a confidence-based orchestration that maximizes automation in situations where the regulatory risk is minimal and maintains human judgment in situations where the exposure is high. The results of human inspection are, over time, added to training data, enhancing the performance of the model on ambiguous and edge cases.

Trust and auditability are also facilitated by explainability mechanisms. The system exposes salient features that affect predictions, including keywords in the description of items or compatibility with familiar category patterns. Such signals enable the compliance experts and auditors to interpret, justify, and justify automated decisions.

5. Validation, Monitoring, and Continuous Learning

Learning-based systems are vital in a system where compliance is necessary and, hence, should be well-tested. There are numerous levels of validation of the architecture, like shadow testing, where the new models are concurrently tested with production systems, as well as historical backtesting on audited datasets. These methods allow the performance evaluation without influence on live tax decisions.

At the transaction level, the monitoring is done to follow the prediction distributions, drift in confidence, and error patterns to identify new risks or regulatory changes. When the metrics are out of the anticipated limits, alerts are issued, which allows proactive intervention.

The system uses the closed-loop learning and validation pipeline, as shown in Figure 3. The result of production classification is constantly checked on drift and anomalies, and low-confidence cases, audit results, and expert reviews are high-quality feedback signals. These signals drive controlled retraining processes whereby new models are tested by shadow deployments and then promoted to production, so that there is continuous improvement without sacrificing regulatory stability. Constant learning pipelines combine human review feedback, audit results, and regulatory changes. To avoid catastrophic forgetting, incremental retraining strategies are used to allow adaptation to new products and changes in policies.

Learning-based systems are critical in environments that require compliance and, therefore, must be robustly validated. The architecture has many layers of validation, such as shadow testing, whereby new models are executed in parallel with production systems and historical backtesting against audited datasets. These techniques permit the performance assessment without affecting live tax decisions.

Constant learning pipelines are a combination of human review feedback, audit outcomes, and regulatory modifications. Incremental retraining strategies are employed to prevent catastrophic forgetting, so that adaptation to new products and new policies can be made.

6. Discussion and Broader Implications

The architecture above demonstrates that machine learning can be successfully and responsibly deployed in global tax compliance when the system is a controlled and open system. With the replacement of manual rule authoring with adaptive classification services, organizations can receive a lot in the sphere of efficiency, accuracy, and scalability without losing regulatory control.

The described principles, such as confidence-aware automation, human-in-the-loop governance, and continuous validation, are generalizable to other compliance-focused areas, such as customs classification, trade controls, and financial reporting.

Table 1: *Key Challenges in Global Tax Classification*

Dimension	Description
Data Quality	Noisy, sparse, and inconsistent item metadata
Product Diversity	Long-tail catalogs with frequent new items
Regulatory Variability	Jurisdiction-specific tax rules and exemptions
Digital Goods	Ambiguous classification across goods/services/IP
Risk Profile	High penalties and audit exposure
System Requirement	Accuracy, explainability, and scalability

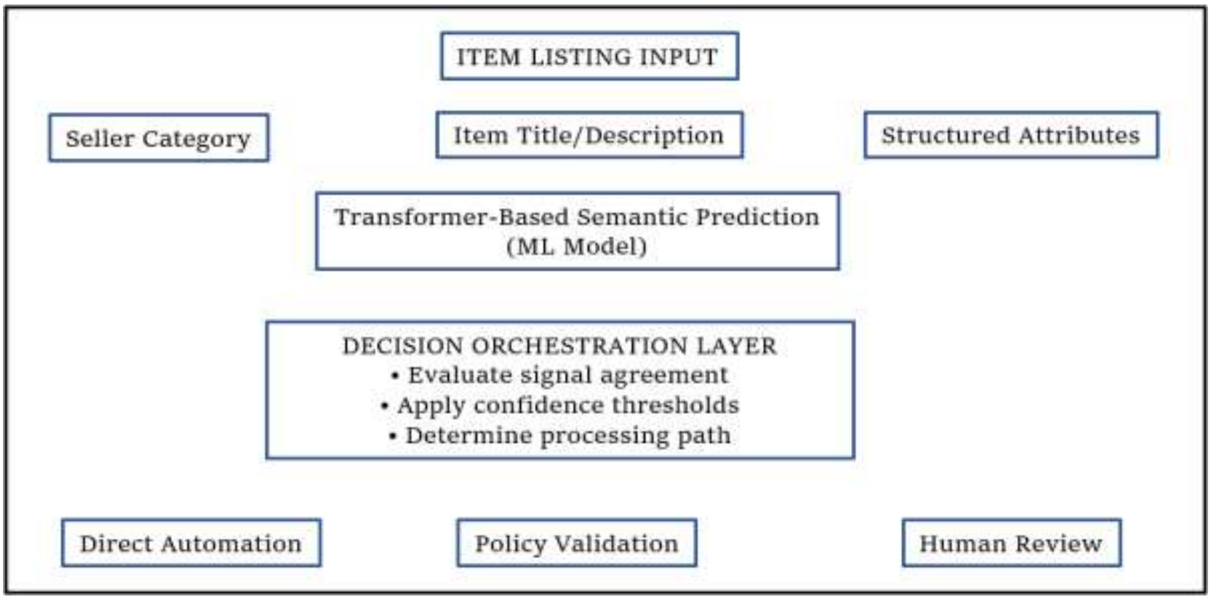

Figure 1: *Multi-Input Item-to-Tax Classification Flow*

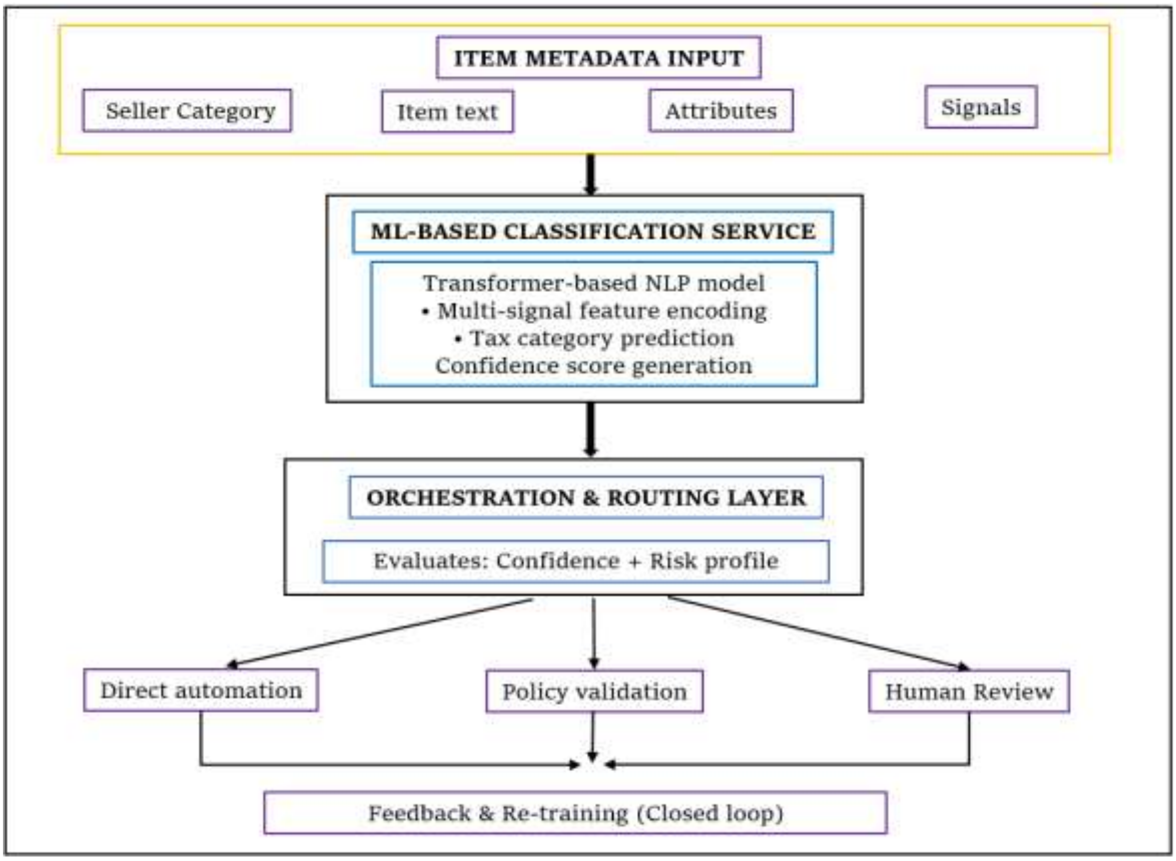

Figure 2: *Multi-Input Item-to-Tax Classification Flow*

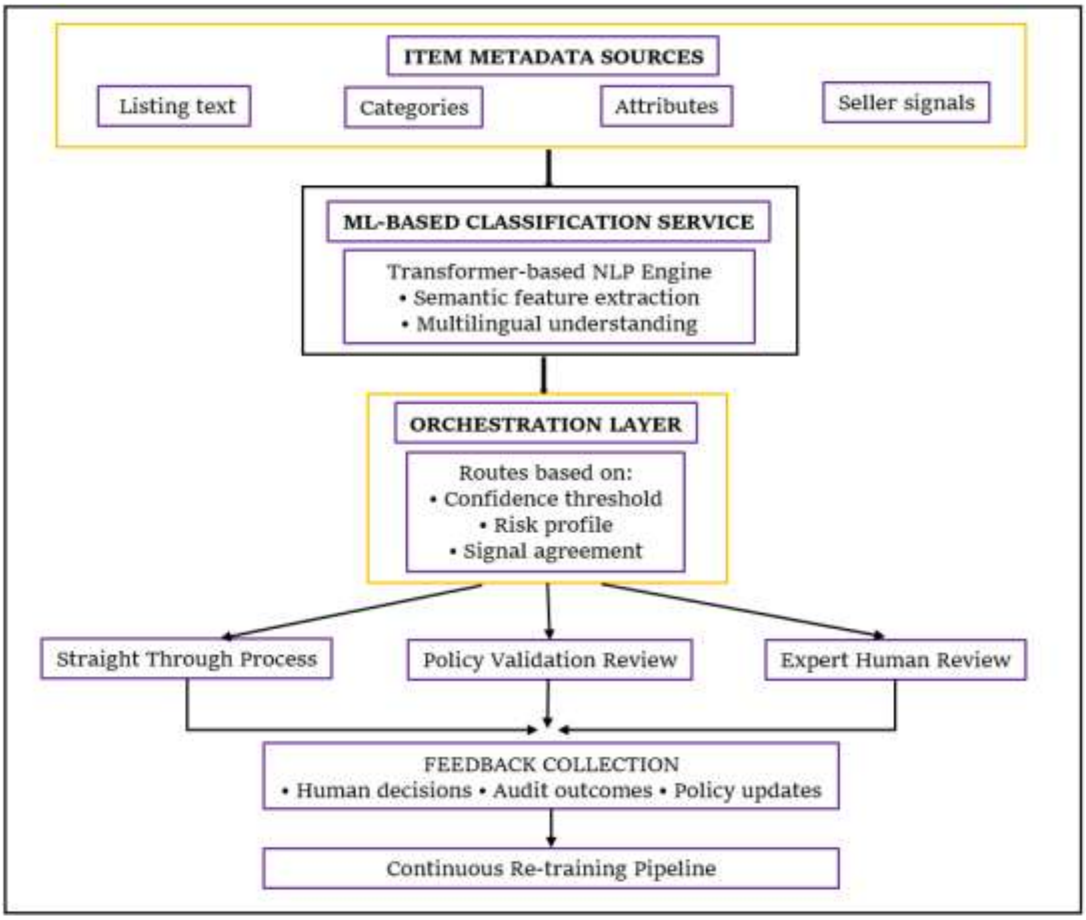

Figure 3: *High-Level Architecture of Intelligent Tax Automation*

Table 2: *Decision Routing Based on Confidence Levels*

Confidence Range	Processing Path	Risk Profile
High	Fully automated	Low
Medium	Policy validation	Moderate
Low	Human expert scrutiny	High

Table 3: *Validation and Learning Mechanisms*

Mechanism	Purpose
Shadow Testing	Safe evaluation of new models
Backtesting	Historical accuracy validation
Monitoring	Drift and anomaly detection
Human Feedback	High-quality training labels
Incremental Learning	Controlled model evolution

7. Conclusions

This article introduces a transformer-based architecture of smart tax automation that directly overcomes the drawbacks of the traditional rule-based and ensemble-driven classification systems. The proposed framework will allow the classification of tax in global marketplaces to be accurate, scalable, and auditable by adding semantic item understanding, discrepancy-conscious decision governance, and confidence-calibrated automation. One of the most important contributions of this work is the fact that high-stakes compliance areas can effectively implement modern machine learning if systems are designed in such a way that transparency, human control, and ongoing validation are explicitly ensured. Instead of viewing automation as a binary decision, the framework makes selective automation operational, i.e., to maximize efficiency, regulatory control, and audit readiness. Beyond tax compliance, the architectural principles described in this article can be used on a broad spectrum of regulated decision-making problems, including custom classification, trade compliance, financial reporting, and risk assessment. These AI systems that are software-defined and governance-aware will prove essential in maintaining confidence, precision, and regulatory obedience as the digital platforms continue to grow around the globe. The article offers a practical foundation of the upcoming academic research involving the domain of machine learning, compliance engineering, and large-scale decision systems, and demonstrates how transformer-based frameworks could be responsibly deployed in the field where accuracy, responsibility, and explainability are paramount.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.

- **Author contributions:** The authors declare that they have equal right on this paper.

- **Funding information:** The authors declare that there is no funding to be acknowledged.

- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.
- **Use of AI Tools:** The author(s) declare that no generative AI or AI-assisted technologies were used in the writing process of this manuscript.

References

[1] Alex Mengden and Andrea Nieder, "International Tax Competitiveness Index 2025", Tax Foundation, 20th Oct. 2025. Available: https://taxfoundation.org/research/all/global/2025-international-tax-competitiveness-index/

[2] Manabu Nose et al., "Leveraging Digital Technologies in Boosting Tax Collection," IMF, May 2025. Available: https://www.imf.org/en/-/media/files/publications/wp/2025/english/wpiea2025089-print-pdf.pdf

[3] Sina Gholamian et al., "LLM-Based Robust Product Classification in Commerce and Compliance," arXiv, 2024. Available: https://arxiv.org/pdf/2408.05874

[4] R. Deepa Lakshmi and N.Radha, "Machine Learning Approach for Taxation Analysis using Classification Techniques", International Journal of Computer Applications, 2011. Available: https://www.ijcaonline.org/volume12/number10/pxc3872322.pdf

[5] Seokho Kang et al., "Multi-class classification via heterogeneous ensemble of one-class classifiers",

ScienceDirect, 2015. Available: https://www.sciencedirect.com/science/article/abs/pii/S0952197615000846
[6] Kushagra Mishra et al., "A hybrid rule-based NLP and machine learning approach for PII detection and anonymization in financial documents", Nature, Jul. 2025. Available: https://www.nature.com/articles/s41598-025-04971-9
[7] Zongjing Cao et al., "Deep Neural Network Confidence Calibration from Stochastic Weight Averaging," MDPI, 2024. Available: https://www.mdpi.com/2079-9292/13/3/503
[8] Bhargavi Tanneru, "Adaptive Compliance: Automating Dynamic Security Controls in Real-Time", International Journal of Multidisciplinary Research and Growth Evaluation, Apr. 2025. Available: https://www.allmultidisciplinaryjournal.com/uploads/archives/20250523163626_MGE-2025-3-043.1.pdf
[9] Wim J. van der Linden, "Review of the shadow-test approach to adaptive testing", Springer Nature, 2021. Available: https://link.springer.com/article/10.1007/s41237-021-00150-y
[10] Vikram Singh, "Adaptive Financial Regulation Through Multi-Policy Analysis using Machine Learning Techniques", IJERT, Apr. 2025. Available: https://www.ijert.org/adaptive-financial-regulation-through-multi-policy-analysis-using-machine-learning-techniques