Asha Judi VARGHEESE, PhD Candidate (corresponding author)
ashajudi81@gmail.com
Sathyabama Institute of Science and Technology (Deemed to be University), Chennai, India

Thankappan SASIKALA, PhD
sasik.cse2008@gmail.com
Sathyabama Institute of Science and Technology (Deemed to be University), Chennai, India

Compound Suitability Prediction in Blockchain-Based Cross-Domain Supply Chain using AdamW5 Optimized FT-Transformer

Abstract. *The safe and effective use of chemical compounds across agriculture, pharmaceuticals, and food industries is vital for public health and environmental sustainability. However, current systems struggle with predicting compound suitability, limiting traceability and data integrity. This study presents an integrated AI-driven blockchain framework to predict the suitability of chemical compounds across these sectors, ensuring transparent traceability in supply chains. The research develops a robust multi-label classification model using an optimised FT-Transformer, which effectively handles complex molecular descriptor data. To improve classification performance, the study introduces the AdamW5 optimiser, which integrates multiple optimisation strategies to enhance convergence stability and generalisation performance. The suitability predictions are securely stored on a Hyperledger Fabric blockchain, providing immutable records, automated access control, and verifiable audit trails. This framework achieved a classification accuracy of 98.45%, with a precision of 98.17% and an F1-score of 97.95%, significantly outperforming baseline models, while ensuring secure and tamper-proof data storage.*

Keywords: *AdamW5 Optimiser, blockchain, compound suitability prediction, cross-domain supply chain, FT-Transformer, smart contract.*

JEL Classification: O0, O1, O3.

Received: 14 June 2025	Revised: 14 June 2025	Accepted: 10 September 2025

1. Introduction

Chemical compounds play a crucial role in various industries, spanning from agriculture to pharmaceuticals and food production. The effective and safe utilisation of chemical compounds plays a key role in agriculture, food processing, and drug development (Salthammer 2024; Hegde et al. 2024). Ensuring the appropriate usage of these compounds is essential for safety, efficacy, and regulatory compliance. In the past, identifying a compound's suitability for a particular domain has relied on human assessments and static databases, which fail to accurately reflect the nuanced

DOI: 10.24818/18423264/59.3.25.08
© 2024 The Authors. Published by Editura ASE. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

needs of each sector (Alaca et al. 3034). Monitoring the creation of compounds and effective management within the food, agricultural, and pharmaceutical supply chains is essential for ensuring product safety (Payandeh et al. 2024). The increased concerns about safety and contamination hazards have refocused its attention on improved supply chain traceability (Xia et al. 2023). Simultaneously, supply chain globalisation has presented difficulties in tracking the source and its organisation of chemical compounds (Ghadge et al. 2023). Occurrences of counterfeit substances and lapses in quality control emphasise the need for robust traceability mechanisms. Blockchain technology has emerged as a promising solution, offering immutable and transparent records that enhance trust across supply chain stakeholders (Ahmed et al. 2023).

Recent technological advancements using blockchain can offer a significant and realistic solution that ensures transparency and eliminates the necessity for a trustworthy centralised authority. The management of agrochemical packaging and supply chain is essential due to its hazardous nature. Many intelligent models have been developed to track agrochemicals with trustworthy and immutable data using blockchain traceability (Monteiro et al. 2024). By employing blockchain technology, stakeholders can ensure end-to-end traceability throughout the food supply chain, facilitating real-time monitoring of critical processes such as sourcing of raw materials, production, storage, distribution, and retail, thereby promoting food safety, quality assurance, and consumer trust (Rashed et al. 2025). In the pharmaceutical supply chain, a secure Barcode system was designed to address tracking and tracing issues, focusing on several key factors such as system reliability, security, implementation cost, and flexibility (Bapatla et al. 2024). Existing chemical information systems often focus on static compound data, which restricts their efficiency in smartly classifying and analysing the suitability of compounds across several areas. There is a significant gap in utilising AI technology to automate and improve the prediction of a compound's suitability based on its unique chemical properties. Furthermore, current systems often encounter issues with traceability, decentralisation, and data tampering, which can impact safety assessments, regulatory compliance, and public trust. To address these issues, the proposed approach designs a robust AI-based multi-label classification model that accurately predicts the suitability of chemical compounds and utilises a blockchain-based traceability system to store these predictions securely. This improves transparency, ensures tamper-proof registration, and creates verifiable audit trails for all stakeholders in the chemical supply chain, thereby enhancing safety and trust in the sector.

1.1 Research Contribution

The primary contribution of the proposed AI-based suitability prediction framework is summarised as follows:

- The proposed work employs the FT-Transformer (Feature Tokenizer + Transformer) model, optimised using an AdamW5 optimiser. It effectively handles

high-dimensional tabular molecular descriptors for multi-label classification of chemical compounds in the agrochemical, pharmaceutical, and food industries.

- We have developed a novel AdamW5 optimiser by integrating AdamW, gradient clipping, gradient centralisation, lookahead, learning rate warm-up, and learning rate schedule to improve convergence speed, training stability, and model generalisation.
- The proposed system combines AI-based suitability prediction with a blockchain-based traceability system, providing tamper-proof, verifiable, and transparent storage of high-confidence compound classification through smart contracts.
- The system incorporates automated access control and compound registration logic via smart contracts, facilitating secure data sharing and providing regulatory audit trails of chemical utilisation throughout supply chains.

The article's remaining content is structured as follows: Section 2 analyses the relevant research on agriculture, food, and pharmaceutical supply chains. The materials and methods are covered in Section 3, which also describes the dataset used, preprocessing methods, and model architecture. Section 4 exhibits the outcomes, along with a comparative analysis and the main findings of the proposed framework. Section 5 concludes the study by condensing the main outcomes.

2. Related Work

In this section, we review and highlight related work found in the literature on AI-based approaches for predicting chemical compound properties and suitability, as well as blockchain applications in agriculture, food, and pharmaceutical supply chains.

2.1 AI-Based Compound Suitability and Property Prediction

(Galushka et al. 2021) presented a deep-learning model for predicting chemical compound properties, with a focus on LogD and binding affinity. Using a variational autoencoder, it generates fingerprints for regression and classification models, demonstrating high accuracy. While it enhances virtual screening efficiency, challenges such as data sparsity and reliance on structural definitions persist. To overcome these limitations, (Kong et al. 2022) developed a new graph neural network, specifically a reduced graph message-passing neural network, for chemical property prediction that integrates pharmacophore information. It demonstrated that pharmacophore-based reduced graph pooling improves prediction accuracy. However, it could not be applicable to all molecular structures.

(Moshkov et al. 2023) aimed to improve drug discovery by assessing chemical structures, imaging, and gene-expression profiles to predict compound bioactivity. By integrating these data, the prediction achieves a higher success rate than using a single modality. This fusion enhances virtual screening efficiency, but it faces challenges such as data sparsity and limited fusion strategies. Similarly, (Abdallah

et al. 2024) develop AI models to predict skin permeability for transdermal drug delivery using molecular descriptors. It utilised machine learning models, including Random Forest (RF), Artificial Neural Network (ANN), Light Gradient Boosting Machine (LGBM), CatBoost, and XGBoost, for analysis. Cluster analysis of FDA-approved drugs revealed distinct permeability patterns, which aid in drug formulation. However, the model faces challenges in predicting high-molecular-weight drugs.

2.2 Agri-food and Pharmaceuticals supply chain

(Bhatia et al. 2023) developed an explainable AI-based faster region with a CNN model for a user-friendly website based on a QR code to evaluate the content of food items. Additionally, a blockchain-based elliptic curve integrated encryption technique was employed to verify this. Even though the proposed model improves the food traceability scenario, implementing these solutions in real-time applications remained difficult. (Sun et al. 2025) created anomalous data processing methods using machine learning to create a blockchain-based traceability model for grain and oil food supply chains to improve the flexibility and validity of data. The model successfully provides comprehensive outlier detection and ensures authenticity and reliability. However, integrating blockchain with the complex structure of the grain and oil supply chain could provide challenges.

(Mishra et al. 2024) developed a blockchain-based framework for online pharma production that eliminates intermediates with hospitals and pharmacies using Ethereum smart contracts to track transactions. This offers an immutable, tamper-proof, and reliable system. The model could enable decentralised applications for processing the transaction to improve scalability. To increase scalability and efficiency, (Abdallah et al. 2023) created distributed ledger technology, which provides a transparent pharmaceutical supply chain and smart contracts. The model enhances data privacy and security. This could be further extended by incorporating a hardware security mechanism that uses a physical unclonable function to authenticate the data.

As evident from these related works, maintaining the safe and effective use of chemical substances across the agricultural, food, and pharmaceutical sectors is becoming increasingly problematic due to limited transparency, manual verification processes, and the absence of real-time traceability in current systems. While various AI-based models have been proposed for predicting chemical properties, many suffer from limitations, including domain-specific focus, poor generalisation to complex or high-molecular-weight compounds, and a lack of integration with traceability mechanisms. With higher risks of regulatory non-compliance, contamination, and abuse, there is a need for an intelligent, automated, and robust system that can accurately predict compound suitability and provide tamper-proof traceability across the supply chain. To address these issues, this research creates an AI-driven framework to accurately forecast the domain suitability of agrochemical, pharmaceutical, or food chemical compounds based on their molecular descriptors.

It uses blockchain-based smart contracts to secure registration and track these predictions to enable transparency, regulatory compliance, and trust across domains for the use of chemicals throughout the supply chain.

3. A Blockchain-Based Framework for Domain-Specific Prediction Traceability

This work presents a blockchain-enabled traceability framework for secure and domain-specific prediction of a chemical compound's suitability, as depicted in Figure 1. The system is constructed with a robust AI-based model using the FT-Transformer architecture, which is tailored for processing high-dimensional tabular data, such as molecular descriptors. The training is optimised using an AdamW5 optimiser that combines gradient clipping, gradient centralisation, lookahead, learning rate warm-up, and learning rate scheduling for rapid convergence, stability, and good generalisation. The model utilises multi-label classification to determine whether a compound is suitable for use in agrochemicals, pharmaceuticals, and/or food applications. Once predictions are generated, the blockchain registration only considers high-confidence predictions. These forecasts, along with compound identifiers, confidence scores, and predicted domain labels, are inserted into a smart contract and immutably stored on a blockchain platform via Hyperledger Fabric. The smart contract not only controls data storage but also implements automated access control, ensuring that all stakeholders can access verifiable and tamper-proof audit trails. This integrated solution ensures that each cross-domain-specific prediction is securely logged and accessible, fostering trust, compliance, and accountability across cross-domain chemical supply chains.

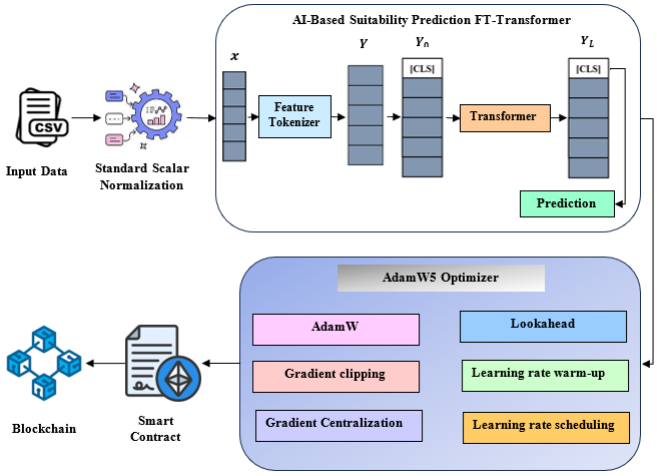

Figure 1. AI-based blockchain-enabled traceability framework for secure and domain-specific prediction of a chemical compound's suitability
Source: Authors' own creation.

3.1 Dataset Description

The PubChemLite Compound Collection for Exposomics is a comprehensive compilation of over 371,000 chemicals from a diverse range of areas and application domains. This invaluable library provides data on molecular structure and composition, annotation categories, and chemical functionality, as well as useful information about associated disorders and diseases. It encompasses fields ranging from numerology to drug discovery, nutrition to toxicology, all enriched with PubMed papers and patents related to each substance. Moreover, the collection includes safety information regarding the pharmacological effects of each compound as well its toxicity profile when exposed in vitro or when metabolised by the liver. For food-related substances, the food-related field provides further details on whether their use is suitable for human consumption or not. With its comprehensive range of annotation, the PubChemLite_31Oct2020_exposomics.csv file provides detailed annotation categories of 371,663 chemicals across a range of applications for research and development. This collection can provide invaluable insight into how the environment affects human health. Compounds for Studying Environmental Exposures Dataset: https://www.kaggle.com/datasets/thedevastator/pubchemlite-compound-collection-for-exposomics-3?resource=download.

3.2 Normalisation using Standard Scalar

The proposed model utilises a standard scalar (Hnamte and Hussain 2023), which is used to standardise the values of numerical features. It changes the data so that the mean becomes 0 and the standard deviation becomes 1. Equation (1) illustrates this process by subtracting the data mean and dividing it by the standard deviation. This centres the data around zero and standardises variability.

$$Z_{scaled} = \frac{x-\mu}{sd} \quad (1)$$

Where x represents the original value, μ represents the mean of the feature, and sd represents the standard deviation.

3.3 AI-based Suitability Prediction using FT-Transformer

Once the data is pre-processed, the next step is to predict domain-specific suitability using an AI model that can learn complex relationships between molecular descriptors, enabling accurate multi-label classification across agrochemical, pharmaceutical, and food domains. FT-Transformer (Wang et al. 2025) is an advanced neural network designed for handling tabular data. It uses a robust self-attention mechanism in the transformer to learn intricate correlations between the features. It is particularly effective in situations where traditional ML models perform poorly because of high dimensionality or the requirements to capture long-range dependencies among the features.

Initially, the FT-Transformer converts the input characteristics into an appropriate representation. Each feature in the input vector a_i is incorporated into a

high-dimensional space using a learnable embedding matrix E as shown in Equation (2).

$$d_i = \mathrm{E} \cdot a_i \tag{2}$$

Where the embedded representation of ith feature is represented by d_i. These embedded features are tokenised to generate a series of feature tokens as represented in Equation (3). This enables the model to capture interactions between diverse features by allowing the transformer to treat each feature distinctly.

$$\mathrm{T} = \{d_1, d_2, \dots, d_k\} \tag{3}$$

Then, the self-attention mechanism allows the model to determine how important each feature token is in relation. The attention mechanism computes a weighted sum of the input tokens using the weights based on token similarity. The mathematical formulation is expressed in Equation (4).

$$Attention\ (\mathcal{Q}, \mathcal{K}, \mathcal{V}) = softmax\left(\frac{\mathcal{Q}\mathcal{K}^T}{\sqrt{e_k}}\right)\mathcal{V} \tag{4}$$

Where $\mathcal{Q}$ is the query, $\mathcal{K}$ is the key, and $\mathcal{V}$ is the value of a linear transformation of the input tokens T. The FT-transformer employs multi-head attention, which combines the output of several self-attention mechanisms in parallel, to improve the model's ability to concentrate on a several input. After the attention mechanism, the tokens are passed through a position-wise feed-forward network that consists of two linear transformations with a ReLU activation function as shown in Equation (5).

$$FFN(\mathrm{T}) = ReLU(\mathrm{T}\mathcal{W}_1 + \mathscr{b}_1)\mathcal{W}_2 + \mathscr{b}_2 \tag{5}$$

Where the learnable parameters are denoted by $\mathcal{W}_1$, $\mathcal{W}_1$, $\mathscr{b}_1$, and $\mathscr{b}_2$. Finally, the output of the feed-forward network is passed through a linear layer to generate the final prediction as expressed in Equation (6).

$$\hat{f} = \mathcal{W}_{out}\mathrm{T} + \mathscr{b}_{out} \tag{6}$$

Where $\mathcal{W}_{out}$ and $\mathscr{b}_{out}$ are the weight and bias of the output layer, respectively.

3.3.1 AdamW5 Optimiser

To further improve the performance and stability of the FT-Transformer, an AdamW5 optimiser is developed. It is an extremely complex and reliable algorithm that aims to integrate and synergise existing optimisers to produce a more effective optimiser. The components involved in this algorithm are AdamW, gradient clipping, gradient centralisation, lookahead, learning rate warm-up, and learning rate scheduling.

The first component of the algorithm is AdamW which includes decoupled weight decay. The next part of the algorithm involves implementing gradient clipping. This sets a threshold for the gradient to the loss function. In this AdamW5, we use the adaptive gradient clipping function, which clips the gradients according to unit-wise ratios of the gradient norms to parameter norms, as expressed in Equation (7). It has less training time and does not require a significant amount of hyperparameter tuning.

$$G_y^t = \begin{cases} \tau \frac{max(|\theta_y^t|, e)}{|G_y^t|} G_y^t & if \frac{|G_y^t|}{max(|\theta_y^t|, e)} > \tau, \\ G_y^t & otherwise. \end{cases} \tag{7}$$

Where G_y^t represents the gradient of the parameter y at time t, θ_y^t represents the model parameter y at the time t, τ represents a clipping threshold, and e is the constant. The next portion of the algorithm is gradient centralisation, regularises the loss function to smooth out the optimisation landscape. It is done by normalising the gradients by deducting their mean before adding them to the optimiser. It smooths out the curve and enhances training speed convergence. This is mathematically expressed as shown in Equation (8).

$$G_{centralized} = \nabla C_y(\theta_{y-1}) - mean\left(\nabla C_y(\theta_{y-1})\right) \quad (8)$$

Where $\nabla C_y(\theta_{y-1})$ represents the gradient of the loss function parameter. Ranger's first optimiser was built using the rectified Adam optimiser and Lookahead. While designing an AdamW5, we used the benefit of Lookahead in the algorithm. As demonstrated in Equation (9), it iteratively updates two sets of weights by selecting a route that looks forward at the series of fast weights produced by another optimiser.

$$\mu_{y+x} = \mu_y + \varepsilon \cdot (\theta_{y+x} - \mu_y) \quad (9)$$

Where μ_y represents the slow weight for the parameter y, θ_{y+x} represents fast weight at x steps, and ε represents the lookahead factor. The next part of the algorithm is learning rate warm-up, which is the most effective regularisation strategy for network training. The learning rate, which is set to 2000 iterations by default for the hyperparameter, is subjected to linear warmup iterations. However, it produces a warm-up that is too lengthy for shorter training runs, as shown in Equation (10).

$$\omega_t = min\left(1, max\left(\frac{1-\varphi_2}{2}.Y, \frac{Y}{y_{warmup}}\right)\right)\eta \quad (10)$$

Where ω_t represents the learning rate at the time t during warm-up, φ_2 represents coefficient, Y represents training step count, y_{warmup} represents a total number of warmup steps, and η represents learning rate. Finally, the last part of the algorithm is the learning rate schedule. This schedule establishes a method for varying the learning rate during training, generally to trade off rapid early learning against precise convergence at later times. In this work, we adopt a cosine annealing schedule, which gradually reduces the learning rate from an initial value to a predefined minimum using a cosine function. This smooth decay helps avoid sharp drops and improves convergence stability. The learning rate at step y is computed as shown in Equation (11).

$$\omega_t = \omega_{min} + \frac{1}{2}(\omega_0 - \omega_{min})\left(1 + cos\left(\frac{y\pi}{Y_{max}}\right)\right) \quad (11)$$

Where ω_0 represents the initial learning rate, ω_{min} represents the minimum learning rate, and Y_{max} represents the total number of training epochs.

3.4 Blockchain-based Traceability System for Secure Storage

After high-confidence predictions are received, it is crucial to ensure that the results are not only securely recorded but also transparently accessible. For this purpose, a blockchain-based system based on Hyperledger Fabric is used for storing the compound predictions immutably. This ensures data integrity, enhances traceability, and ensures that all stakeholders can verify predictions in a tamper-proof manner.

The proposed blockchain-based system for domain appropriateness across agriculture, food, and pharmaceutical products aims to trace, share, and register specific data throughout the entire supply chain. The system uses this to provide transparency, tamper-resistant registration, and auditable trails for all the stakeholders across the chemical supply chain.

3.4.1 Smart Contract

To automate and secure the process of prediction logging and data access on the blockchain, a smart contract mechanism is employed. A smart contract is a form of a computer program that executes on the blockchain and can automatically execute when specific conditions are satisfied. The proposed work utilises solidity language to write smart contracts. Through this, the system records the traceability information of agriculture, food, and pharmaceutical products, tracks the status of these products, and manages the supply chain. This automation promotes regulatory compliance and cross-domain trust.

A smart contract (Bandhu et al. 2023) is implemented to automate the secure logging of prediction results. It records only high-recommendable compounds, each entry is timestamp, tamper-proof, and verifiable, and only authorised users can access the system. Each recorded prediction entry is cryptographically linked to its metadata and stored immutably on the blockchain. This eliminates the risk of tampering and ensures auditability throughout the system.

4. Results and Discussion

Table 1. Hyperparameters of AdamW5 optimiser

Parameter	Values
Learning Rate	0.001
weight_decay	0.01
clip_grad_norm	1.0
gradient_centralization	True
use_lookahead	True
warmup_steps	1000
lr_schedule	cosine
total_steps	10000

Source: Authors' processing.

In this section, the proposed AI-based suitability prediction framework was analysed to evaluate its performance across several metrics. This has been executed on a Windows 10 system. TensorFlow version 2.17 and Python 3.12 were used to create the DL models. The hardware specifications include an Intel Core i5 6500 CPU running at 3.20 GHz, 8.0 GB of DDR3 RAM, and an Intel(R) HD Graphics 530 for graphics. Hyperparameter settings utilised in the AdamW5 optimiser are detailed in Table 1.

4.1 Performance Evaluation of the Proposed Framework

This section examines the proposed system's performance in effectively predicting the compounds based on the molecular description, with an emphasis on its classification capabilities. The proposed model performs exceptionally well on several metrics, including accuracy, precision, F1-score, recall, Matthews Correlation Coefficient (MCC), and specificity at 98.45%, 98.17%, 99.68%, 99.40%, 98.35%, and 98.90%, as shown in Table 2. This demonstrates how well the proposed model accurately predicts the suitability of compounds. Additionally, the model exhibits a low propensity for mistakes, with few false positives and false negatives, as evidenced by its False Positive Rate (FPR) of 0.016 and False Negative Rate (FNR) of 0.23.

Table 2. The Proposed Framework's Performance Metrics

Metrics	Accuracy (%)	Precision (%)	F1-score (%)	Recall (%)	MCC (%)	Specificity (%)	FPR (%)	FNR (%)
Values	98.45	98.17	99.68	99.40	98.35	98.90	0.016	0.23

Source: Authors' processing.

4.2 Blockchain Analysis

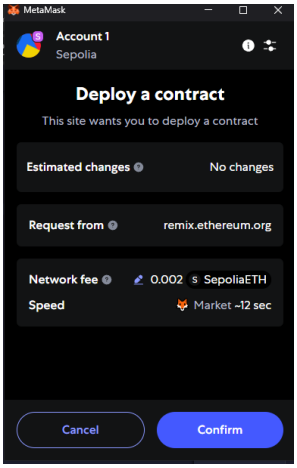

Figure 2. MetaMask Smart Contract Deployment
Source: Authors' own creation.

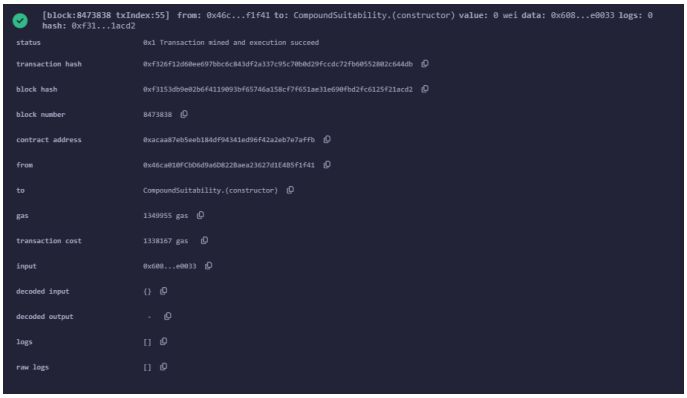

Figure 3. Smart Contract Deployment Confirmation on Sepolia
Source: Authors' own creation.

Figure 4. Blockchain-Logged Cross-Domain Chemical Usage Dashboard
Source: Authors' own creation.

This section presented an end-to-end deployment and logging of domain-specific compound suitability prediction results on a blockchain using the Sepolia Ethereum testnet. Figure 2 shows the MetaMask interface prompting the user to deploy a smart contract from Remix IDE, including details such as the request origin, estimated gas fees, and transaction confirmation time. Figure 3 confirms successful contract deployment on the Sepolia blockchain, including transaction hash, sender and receiver addresses, gas usage, and timestamp. Figure 4 displays the front end of the blockchain-logging dashboard, listing each image alongside its compound ID, prediction suitability across agriculture, food, and pharmaceutical, model version, timestamp, and cryptographic hash. This secure, transparent, and verifiable logging mechanism enhances trust and traceability in cross-domain chemical usage across the supply chain.

Figure 5 shows the exchange between the threshold for prediction confidence and blockchain transaction behaviour. The higher the threshold, the fewer compound predictions are selected for blockchain logging, saving aggregate gas costs. However, this comes at the cost of excluding potentially suitable compounds. This balance is critical in ensuring both cost-efficiency and trustworthy traceability in real-world supply chain deployments.

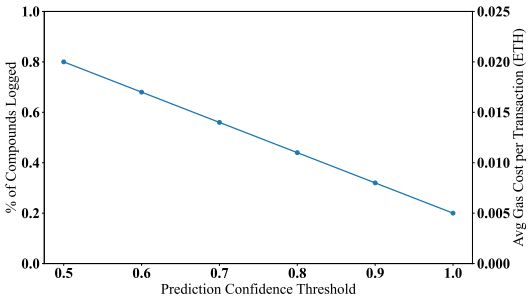

Figure 5. Prediction Confidence vs. Blockchain Logging Cost
Source: Authors' own creation.

Table 3 shows the comparison of the standard and AdamW5 optimisers in terms of the percentage of high-confidence predictions (confidence ≥ 0.90) and the number of predictions stored on the blockchain. In comparison to the standard AdamW optimiser, the enhanced version resulted in a 14.3% improvement in the percentage of high-confidence predictions from 74.8% to 89.1%. This enhancement directly contributed to the sum of predictions that qualify for blockchain logging, leading to a larger number of reliable and traceable entries in the supply chain ledger.

Table 3. Impact of Optimiser on High-Confidence Prediction Logging

Optimiser	High-confidence Rate (%)	Logged Predictions
Standard AdamW	74.8	748
AdamW5 (Proposed)	89.1	891

Source: Authors' processing.

4.3 Ablation Study

In this section, we analyse the results of the ablation studies. Table 4 shows the results of the proposed AdamW5 optimiser with different epochs and learning rates.

To validate the proposed model's performance, we conducted an ablation study by contrasting its performance with various optimisers. This outcome demonstrates how the proposed AdamW5 optimiser can improve the model's performance in terms of accuracy, precision, recall, and F1-score among other measures.

Table 4. Ablation Analysis of the Proposed Framework

Optimiser	Epochs	Learning Rate	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Adam	5	0.0005	95.25	94.90	95.67	95.30
	10		95.50	95.20	95.90	95.55
	15		95.75	95.44	96.17	95.80
	20		96	95.70	96.40	96.05
AdamW	5	0.0008	95.75	95.84	96.04	95.80

Optimiser	Epochs	Learning Rate	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
	10		96	95.70	96.40	96.05
	15		96.25	95.55	95.81	96.30
	20		96.50	96.20	96.90	96.55
RAdam	5	0.001	96.25	95.97	96.60	96.34
	10		96.50	96.20	96.90	96.55
	15		96.75	96.47	97.13	96.80
	20		97	96.70	97.40	97.05
Ranger	5	0.0007	96.51	96.85	97.09	96.80
	10		97	96.70	97.40	97.05
	15		97.20	96.95	97.69	97.30
	20		97.50	97.42	97.90	97.55
Ranger21	5	0.0009	97.21	96.95	97.68	97.30
	10		97.68	97.12	97.9	97.55
	15		97.72	97.81	98.13	97.80
	20		98	97.70	98.40	98.05
AdamW5 (Proposed)	5	0.001	97.75	97.45	98.05	97.80
	10		98	97.70	98.40	98.05
	15		98.25	97.95	98.65	98.30
	20		98.50	98.17	99.06	97.95

Source: Authors' processing.

4.4 Comparative Analysis with Baseline Models

This section compares the effectiveness of the proposed model with different baseline methods for predicting the compounds' suitability. Table 5 shows the performance of the FT-Transformer model compared with other models, Random Forest, MLP, XGBoost, TabNet, and TabTransformer on the suggested framework.

Table 5. Comparative evaluation of the proposed model with baseline approaches

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Random Forest	94.23	93.8	94.5	94.01
MLP	95.1	94.7	95.2	94.95
XGBoost	96.45	95.61	96.3	96.09
TabNet	96.8	95.85	96.7	96.22
TabTransformer	97.6	96.2	97.41	97.85
FT-Transformer (Proposed)	98.45	98.17	99.06	97.95

Source: Authors' processing.

The FT-Transformer model used in the proposed system achieves the optimal overall performance, with an accuracy of 98.45%, a precision of 98.17%, a recall of

99.06%, and an F1-score of 97.95%. This outstanding performance is due to the ability of the FT-Transformer to handle high-dimensional tabular data through attention-based feature interactions, which makes it suitable for predicting compound suitability across a wide range of domains.

4.5 Discussion

The presented work tackles the critical challenge of determining the cross-domain applicability of chemical molecules between agrochemical, pharmaceutical, and food domains while ensuring secure and tamper-proof traceability across the supply chain. The conventional systems mainly rely on fixed databases and human judgment, which are error-prone, inflexible, and do not offer real-time or decentralised traceability. Additionally, current solutions fail to include intelligent multi-label classification or provide transparent audit trails that are verifiable by supply chain stakeholders. To overcome these limitations, the study proposes an AI-driven multi-label classification model based on the FT-Transformer for precise suitability prediction from high-dimensional molecular descriptor data and a blockchain-based traceability system based on Hyperledger Fabric and smart contracts to securely record and handle the prediction results.

The FT-Transformer model, advanced by an AdamW5 optimiser, effectively extracts intricate relationships between features in the dataset. The optimiser combines multiple components including gradient clipping, centralisation, lookahead updates, learning rate warm-up, and cosine annealing to improve convergence speed, stability, and generalisation. This outcome shows excellent classification performance, as demonstrated by 98.45% accuracy, 98.17% precision, and an F1-score of 97.95%. To validate the framework, a thorough evaluation was conducted through benchmark metrics, heatmap, ROC curves, and ablation studies comparing different optimisation approaches. The AdamW5 optimiser consistently outperformed baseline optimisers like Adam, AdamW, RAdam, Ranger, and Ranger21. Furthermore, a comparative analysis with state-of-the-art models including Random Forest, XGBoost, TabNet, and TabTransformer confirmed the superiority of the FT-Transformer in this domain. In addition to model evaluation, the blockchain component was validated through practical deployment on the Sepolia Ethereum testnet. Smart contracts were designed to automate compound registration, facilitate access control, and offer immutable logs. The system demonstrated efficient transaction times and consistent gas costs, showing feasibility for real-world implementation. The blockchain dashboard also supports end-to-end transparency by displaying prediction metadata, timestamps, and cryptographic hashes for every logged compound. Combining AI and blockchain technologies ensures both high prediction accuracy and verifiable data integrity fulfilling regulations, safety, and trust requirements in cross-domain supply chains.

5. Conclusions

In this study, we introduced an integrated framework that combines an optimised FT-Transformer model with a blockchain-based traceability system to address the problem of predicting chemical compound suitability in agrochemical, pharmaceutical, and food contexts. By using the FT-Transformer attention mechanism and improving its training using an AdamW5 optimiser, the model effectively managed high-dimensional molecular descriptor data and attained a high classification accuracy of 98.45% and generalisation. To ensure safe, transparent, and tamper-proof prediction logging, we utilised a blockchain implementation with Hyperledger Fabric through smart contracts, allowing immutable record-keeping as well as automatic access control. Experimental findings validated the improved performance of the model over baseline techniques, while blockchain deployment demonstrated feasibility for real-world traceability. In future work, we plan to expand the framework by incorporating real-time data streams and integrating additional domain-specific ontologies for improved contextual prediction.

References

[1] Abdallah, R.M., Hisham E.H., Hammad, A. (2024), *Predictive modeling of skin permeability for molecules: Investigating FDA-approved drug permeability with various AI algorithms. PLOS Digital Health,* 3(4), e0000483.

[2] Abdallah, S., Nizamuddin, N. (2023), *Blockchain-based solution for pharma supply chain industry. Computers & Industrial Engineering,* 177, 108997.

[3] Ahmed, W.A., MacCarthy, B.L. (2023), *Blockchain-enabled supply chain traceability–How wide? How deep?. International Journal of Production Economics*, 263, 108963.

[4] Alaca, Y., Emin, B., Akgul, A. (2024), *A comparative study of deep learning models and classification algorithms for chemical compound identification and Tox21 prediction. Computers & Chemical Engineering*, 189, 108805.

[5] Bandhu, K.C., Litoriya, R., Lowanshi, P., Jindal, M., Chouhan, L., Jain, S. (2023), *Making drug supply chain secure traceable and efficient: a Blockchain and smart contract based implementation. Multimedia Tools and Applications*, 82(15), 23541-23568.

[6] Bapatla, A.K., Mohanty, S.P., Kougianos, E. (2024), *Pharmachain 3.0: efficient tracking and tracing of drugs in pharmaceutical supply chain using blockchain integrated product serialization mechanism. SN Computer Science*, 5(1), 149.

[7] Bhatia, Surbhi, Albarrak, A.C. (2023), A *blockchain-driven food supply chain management using QR code and XAI-faster RCNN architecture. Sustainability,* 15(3), 2579.

[8] Galushka, M., Swain, C., Browne, F., Maurice, D. (2021), *Mulvenna, Raymond Bond, and Darren Gray. "Prediction of chemical compounds properties using a deep learning model. Neural Computing and Applications,* 33(20), 13345-13366.

[9] Ghadge, A., Bourlakis, M., Kamble, S., Seuring, S. (2023), *Blockchain implementation in pharmaceutical supply chains: A review and conceptual framework. International Journal of Production Research*, 61(19), 6633-6651.

[10] Hegde, V., Bhat, M.P., Lee, J.H., Kurkuri, M.D., Kim, C.S., Lee, K.H. (2024), *Carbon-based nanomaterials: Multifaceted role in agrochemical recognition, remediation, and release. Nano Today*, 57, 102388.

[11] Hnamte, V., Hussain, J. (2023), *DCNNBiLSTM: An efficient hybrid deep learning-based intrusion detection system. Telematics and Informatics Reports*, 10, 100053.

[12] Kong, Y., Zhao, X., Liu, R., Yang, Z., Yin, H., Zhao, B., Wang, J., Qin, B., Yan, A. (2022), *Integrating concept of pharmacophore with graph neural networks for chemical property prediction and interpretation. Journal of Cheminformatics,* 14(1), 52.

[13] Mishra, R., Dharavath, R., Nazeeruddin, M., Mondal, B. (2024), *Blockchain enabled secure pharmaceutical supply chain framework with traceability: An efficient searchable pharmachain approach. Cluster Computing,* 27(10), 13621-13641.

[14] Monteiro, E.S., Mignoni, M.E., Righi, R.D.R., Costa, C.A.D., Kunst, R., Barbosa, J.L.V., Alberti, A.M. (2024), *On proposing an intelligent model for tracking agrochemicals. Internet Technology Letters*, e382.

[15] Moshkov, N., Becker, T., Yang, K., Horvath, P., Dancik, V., Wagner, B.K., Clemons, P.A., Singh, S., Carpenter, A.E., Caicedo, J.C. (2023), *Predicting compound activity from phenotypic profiles and chemical structures. Nature communications,* 14(1), 1967.

[16] Payandeh, R., Delbari, A., Fardad, F., Helmzadeh, J., Shafiee, S., Ghatari, A.R. (2024), *Unraveling the potential of blockchain technology in enhancing supply chain traceability: A systematic literature review and modeling with ISM. Blockchain: Research and Applications*, 100240.

[17] Rashed, M.S., Fakhry, S., Satour, R., Pathania, S. (2025), *Blockchain Technology in Food Supply Chain Management: Enhancing Traceability, Safety, and Quality. In Food and Industry 5.0: Transforming the Food System for a Sustainable Future* (pp. 427-455), *Cham: Springer Nature Switzerland*.

[18] Salthammer, T., (2024), *Assessment of methods for predicting physical and chemical properties of organic compounds. Indoor Environments*, 100031.

[19] Sun, F., Wang, P., Zhang, Y., Kar, P. (2025), *βFSCM: An enhanced food supply chain management system using hybrid blockchain and recommender systems. Blockchain: Research and Applications* 6(1), 100245.

[20] Wang, Y., Tang, Y., Ma, J., Chen, Z., Cui, C., Li, M., ... Zhang, W. (2025), *Contrastive Learning with Transformer to Predict the Chronicity of Children with Immune Thrombocytopenia. IEEE Journal of Biomedical and Health Informatics*.

[21] Xia, J., Li, H., He, Z. (2023), *The effect of blockchain technology on supply chain collaboration: A case study of lenovo. Systems*, 11(6), 299.