STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **13**, Month 2025, pp 0–16.
Published online in International Academic Press (www.IAPress.org)

A Comprehensive Trust Evaluation Model for Financial Service Providers Using Fuzzy Inference Systems

Alaa R. Madkour [1,*], Mahmoud A. A. Mousa [1,2], Ibrahim Ziedan [1]

[1] *Department of Computer and Systems Engineering, Zagazig University, Egypt*
[2] *Computer Science, Heriot-Watt University, Dubai, UAE*

Abstract This paper presents a hybrid trust evaluation model for financial service providers based on fuzzy inference systems (FIS) and machine learning methods. The proposed model aggregates FSLA compliance measures, operational performance information, and user feedback to calculate dynamic, multidimensional trust scores. The model utilizes both the strengths of fuzzy logic in handling uncertainty and ambiguity, as well as the predictive power and real-time robustness of machine learning. The effectiveness of this hybrid method in overcoming the constraints of existing trust evaluation frameworks was demonstrated by the results, such as their static style, reliance on subjective evaluations, and lack of integration across crucial variables. Moreover, the quantitative evaluation indicated good accuracy, precision, and recall, highlighting the model's reliability and practical application. The suggested framework can evolve into a more versatile and powerful instrument for trust evaluation, thereby enhancing its contributions to the financial industry and beyond.

Keywords fuzzy inference systems, Machine Learning, Trust Evaluation systems, FSLA Compliance, Defect Tracking

DOI: 10.19139/soic-2310-5070-2373

1. Introduction

*T*rust is an essential foundation of the financial business environment, on which service providers' relationships with clients are built. Its impact is at the level of critical decisions, i.e., customer retention, appetite for investment, and overall market stability. However, trust is not a fixed value; trust is dynamically constructed through the reliability, transparency, and ethical behaviors of financial service providers. As the financial landscape shifts toward digitization, trust constructs are becoming more complex owing to the availability of fraud, data breaches, and operational failures. These challenges have also heightened the need for highly robust, multi-parameter trust-assessment models that can be flexible enough to cope with current financial environments [1].

In most traditional trust measurement approaches, subjective evaluations like user reviews or simple reputation scores are used. Although such methods reflect part of trust, they are inherently limited in scope. They omit important considerations, such as regulatory compliance, fraud prevention, and real-time operational performance, which all have a major impact on the trustworthiness of a financial provider [2]. For instance, a service provider may achieve a high user review rating yet lack adherence to core regulatory requirements such as anti-money laundering (AML) and know-your-customer (KYC) frameworks, which could put clients at undue financial risk [3].

The concept of the Financial Service Level Agreement (FSLA) has become one of the most important policy tools for overcoming those gaps. FSLA metrics include regulatory adherence, fraud prevention measures, transaction accuracy, and service availability—parameters that reflect both the legal and operational reliability of

*Correspondence to: Alaa.R.Madkour@gmail.com

ISSN 2310-5070 (online) ISSN 2311-004X (print)

financial providers. Compliance with FSLA standards not only helps meet legal requirements but also strengthens consumer trust by reflecting a provider's dedication to accountability and disclosure. Although critically important, FSLA compliance is frequently neglected in the current trust evaluation models, whose focus is typically either user feedback or a single performance metric [4].

Along with compliance, operational reliability, and fraud prevention are critical factors of trust in financial services. Operating in high stakes, financial institutions are vulnerable to data breaches, system failures, or service interruptions with disastrous effects. These operational hazards demand the creation of trust evaluation systems able to combine several data sources, including operational performance records, defect tracking data, and real-time alarms. Such systems can offer a more complex and flexible evaluation of trust, therefore addressing the dynamic character of financial service delivery [5].

Promising answers to these problems can be found in technologies like machine learning and fuzzy inference systems (FIS). FIS is excellent at managing ambiguity and uncertainty, allowing systems to efficiently handle inaccurate data and subjective input. For instance, [6] proposed a fuzzy model for detecting and predicting cloud service quality violations, which demonstrated the capability of fuzzy systems in proactive handling of performance degradation possibilities in dynamic environments. Similarly, the fuzzy-AHP approach was applied for prioritizing trust criteria in fog computing services, as shown in [7]. It underlines the necessity to integrate multi-criteria decision-making frameworks into trust evaluation systems to make the outcome more accurate and relevant. Meanwhile, machine learning boosts predictive capabilities, allowing trust evaluation systems to recognize trends, foresee errors, and adjust to real-time changes.

By integrating FIS with machine learning, financial institutions can construct a complete trust evaluation model that incorporates different data sources, delivering a dynamic and actionable trust score [8, 9]. An adaptive weighting mechanism is also utilized to address the limitations of traditional, static trust evaluation models in the dynamic financial landscape. This mechanism adjusts the importance of trust parameters—such as FSLA compliance, operational performance, and defect tracking—based on real-time regulatory updates and market trends. It enhances model accuracy and responsiveness to emerging risks and priorities, such as fraud prevention or system reliability during market volatility.

This paper proposes a Trust Evaluation Model combining Fuzzy Inference Systems and machine intelligence methods with an adaptive weighting mechanism to address these gaps. The proposed model has inputs from multi-dimensional data sources including FSLA compliance, operational performance, and user feedback, and thus generates a dynamic, real-time trust score for financial providers. The objectives of our research paper are:

1. Develop a comprehensive trust evaluation model specifically for financial service providers using FISs.
2. Evaluate trust through the integration of several parameters.
3. Implement real-time trust computation so that it can be automatically adjusted to the dynamic regulatory and operational scenarios.
4. Offer trust scores that would correspond to stakeholders who aim to select providers with a good reputation in the financial market.
5. Confirm the model's effectiveness by testing it out through simulated data and the final results.

The remainder of this paper unfolds as follows: Section 2: The literature review provides an exploration of existing trust evaluation models and the used evaluation methods. Section 3: The methodology describes the proposed model and the integration of Fuzzy Inference Systems (FIS) and machine learning in trust evaluation. Section 4: Implementation and Results evaluates the performance of the model and discusses obtaining the results. Section 5: Conclusion summarizes the findings of the research and suggests future research directions.

2. Literature Review

*T*rust evaluation is vital for assuring the trustworthiness, openness, and accountability of service providers, especially in the financial arena. Traditional trust evaluation frameworks generally rely on reputation systems and user feedback mechanisms, which are widely deployed in fields including e-commerce, cloud computing, and

social reviewing systems. A complete taxonomy of reputation systems was given by [10], noting their strengths and drawbacks. While these systems rely on user ratings and previous behavior to determine trust, they fail to account for dynamic and objective performance measures crucial to financial systems.

The contribution of [11] consists of the development of the multilevel trust management model in the service-oriented context, structured as a hierarchical mechanism to gain advantages in terms of scalability and flexibility. Recent developments concern the evidence-based trust evaluation model in cloud services using fuzzy logic that ensures effective integration among subjective-objective metrics towards a more balanced trust measure [12].

From the point of view of IoT service interactions, effective and efficient communications highly depend on the trust and reputation systems of common IoT service users. Battah deligated a decentralized trust and reputation system based on distributed ledger technology and smart contracts [13]. This approach resolves the scalability constraints noted in classic centralized systems by employing decentralized storage and blockchain technologies, thus improving the perception and evaluation of trust in IoT systems.

As far as the domain of trust models is concerned, [14] developed a hybrid trust model for vehicular social networks that reinforces communications trust with social trust parameters. While assessing vehicles, this model considers vehicles, their communication behavior, and the driver's social features, resulting in a comprehensive evaluation of applicable trust within vehicular networks. Results from simulations showed improvements in the accuracy of trust evaluations and the effectiveness of network communications.

In the financial sector, trust must be evaluated across various dimensions, including regulatory compliance, operational reliability, and fraud prevention. [1] emphasizes the relevance of financial innovation in enabling financial inclusion and offers a trust construct that encompasses both subjective user views and measurable system-level reliability. The same [3] analyzes how blockchain technology promotes trust in digital Islamic banking systems by guaranteeing transparency and security, filling major shortcomings in traditional trust models. However, current systems are still missing real-time adaptability and integration of multi-source data, which are crucial for evaluating trust dynamically in financial services.

Fuzzy logic has appeared as a valuable method of coping with the intrinsic uncertainty and ambiguity within trust assessment. In contrast to binary or deterministic approaches, Fuzzy Inference Systems (FIS) permit trust to be measured on a scale encompassing subtleties in subjective information and qualitative data. Cherkassky points to the flexibility of fuzzy logic in decision-making systems, particularly in dealing with inaccurate information [15].

In trust evaluation, FIS offers a means to merge multiple inputs, including performance metrics, compliance records, and user opinion, into a combined trust value. The function of fuzzy rule-based systems in trust assessment has been thoroughly investigated. An introduction of a fuzzy rule-based expert system that evaluates the reliability of cloud service providers through a systematic assessment of several service criteria is represented by [16]. Their work underlines the promise of fuzzy systems in addressing the subjective components of trust while ensuring operational reliability. Moreover, a fuzzy-based trust evaluation framework for fog computing was created by [5], highlighting how such systems can integrate behavioral analysis with performance measurements to give robust trust evaluations in decentralized infrastructures. Similarly, [4] made use of fuzzy logic to implement a 3-layered trust management architecture for mobile edge computing that allows on-the-fly trust evaluations in high uncertainty environments. These works show the scalability and robustness of FIS for trust systems.

Although fuzzy logic has benefits, it is underapplied in financial trust assessment. For example, [17] stresses the necessity for defuzzification approaches in decision-making processes, and in this context, fuzzy logic can be used to perform a critical duty in multi-criteria evaluation. Yet, due to their inability to cover essential features such as regulatory compliance or fraud detection, present fuzzy-based work is often less beneficial in financial applications.

Because machine learning approaches can accurately predict outcomes and evaluate massive, multi-dimensional datasets, they have become increasingly popular in the evaluation of trust. Anomaly detection and trust scoring are common applications for models such as Gradient Boosting Machines (GBM) and other ensemble learning techniques. The efficiency of GBM in predicting trust levels based on previous performance data and user interactions was provided by [8]. These models excel in spotting patterns and trends, making them extremely ideal for dynamic trust evaluation in financial services.

However, a serious challenge in regulatory contexts is the lack of interpretability in most standalone machine-learning approaches where transparency is essential. It is indicated by [18] that the integration of machine learning with fuzzy logic has the potential to bridge the gap by embedding predictive accuracy with explainability. For example, machine learning can predict performance anomalies or compliance breaches, while fuzzy logic may provide an interpretable trust score based on qualitative and quantitative inputs [19].

The integration of machine learning and fuzzy logic facilitates the dynamic modification of trust scores in reaction to real-time fluctuations. A formulation of a multi-criteria trust assessment framework employing subjective logic and machine learning methodologies demonstrates that hybrid approaches can improve the resilience and precision of trust evaluation systems represented by [2].

Operational reliability is an essential factor in trust assessment, especially in financial systems, where the disruption of services or compliance abuse can grossly reduce trust. Techniques for anomaly detection and defect tracking are important means of finding irregularities that could point to fraud or operational failure. Trust-based anomaly detection in emerging sensor networks was used by [20], emphasizing its relevance in recognizing unusual behavior in complex systems. Similarly, a trust evaluation approach for industrial control systems, which integrates anomaly detection and multi-attribute trust rating, has been introduced by [21].

In financial services, defects such as transaction delays, system failures, and compliance violations must be monitored and considered in trust assessments. Incorporating real-time defect tracking tools into trust evaluation systems guarantees that trust scores match the operational reliability of service providers. For example, a fuzzy logic-based security trust evaluation system for IoT contexts was proposed by [9], illustrating the need to integrate defect tracking with performance measurements. The merging of fuzzy logic with machine learning constitutes a substantial development in trust evaluation systems. Fuzzy logic provides a solid framework for addressing uncertainty and qualitative inputs, while machine learning boosts prediction skills and scalability. This hybrid method assures that trust evaluation systems can evaluate multi-source data dynamically and deliver actionable insights.

The viability of hybrid models in fog and mobile edge computing was demonstrated by [4, 5], in which trust is evaluated by performance metrics, behavioral analysis, and system reliability. By combining fuzzy logic and machine learning, the validity and interpretability of the resultant system are guaranteed, and this leads to the solution of the key limitations of traditional approaches. However, the application of such hybrid models for trust assessment in finance is still rather limited, as it reveals a significant gap in the existing research.

Based on the foregoing the gaps in existing research can be listed as follows:

1. Limited Integration of Key Parameters: Current models are not able to include regulatory compliance, operational performance, and fraud detection into a single framework [4, 10].
2. Static trust evaluation: Most systems are not real-time adaptive, and of course, such systems are not good for dynamic financial worlds [2, 3].
3. Insufficient use of hybrid models: The application of fuzzy logic and machine learning has great promise in trust assessment but has been underrepresented [5, 8].
4. Neglect of Defect Tracking: However, for most models, defect tracking and anomaly detection in real-time are not deemed to be an important part of trust assessment [20, 21].

To overcome these gaps, this paper introduces a hybrid trust evaluation model based on fuzzy inference systems and machine learning methods. By combining FSLA compliance, operational measurements, user feedback, and defect tracking, the model offers an interactive real-time trust evaluation environment for the use of financial service institutions.

3. Model Architecture and Methodology

*T*his section presents the architecture and methodology of the trust evaluation model for financial providers based on Fuzzy Inference Systems (FIS). By considering important factors like performance metrics, regulatory compliance, and the tracking of defects and alerts the model assesses the reliability of financial service providers. For the architecture of the model, it is composed of three layers connected together:

3.1. Layer I: Input Data

As shown in Figure 1, There are three main sources of the data used in this model:

1. User Feedback: User reviews and ratings provide insight into how reliable a service provider is perceived to be. it allows users to evaluate their experiences based on various factors such as security and service dependability. To ensure comparability with data from other sources, the feedback is transformed into normalized values.
2. Transaction Logs: These logs include detailed records of the provider's transactions along with information on system uptime processing times, success rates, and error rates. These objective operational measures are essential for assessing the dependability and performance of a provider.
3. Regulatory Reports: Compliance information from Know Your Customer (KYC) frameworks Anti-Money Laundering (AML) and other industry regulations are examples of regulatory data. Legal and ethical standards are crucial in the financial industry and these reports aid in evaluating the provider's compliance with them.

3.2. Layer II: Trust Parameter

The model evaluates trust through three parameters:

Parameter 1: FSLA Compliance Score

FSLA compliance Score is essential for assessing a financial service provider's legal and operational conformance. It is generated from four important metrics:

1. Regulatory Adherence: it is one of the main components used to determine the FSLA Compliance score. In order to maintain financial stability and reduce risks in the financial industry this metric assesses the provider's compliance with AML (Anti-Money Laundering) and KYC (Know Your Customer) laws. Compliance with these regulations assures that providers function within legal frameworks that shield financial institutions and users from illicit activities like fraud and money laundering. The audit findings compliance certificates and penalty records that we use to evaluate regulatory adherence are all normalized for consistency. Providers who meet these crucial requirements are rated as more trustworthy because the FSLA Compliance Score which is largely based on Regulatory Adherence has a direct influence on the final Trust Score.
2. Fraud Prevention: Fraud prevention is a fundamental element of confidence in the financial sector. This measure examines a provider's history of fraud detection and prevention, including the usage of security procedures like multi-factor authentication (MFA), encryption, and other fraud detection mechanisms. The service's ability to protect users' financial transactions and personal data improves with a higher historical fraud prevention rate.
3. Transaction Accuracy: Transaction accuracy is calculated from the transaction logs by measuring the error rate in financial transactions (e.g., transaction failures, and wrong transactions). This score demonstrates how trustworthy the provider's systems are in ensuring correct and successful transactions.
4. Service Availability: This metric assesses the provider's system uptime and availability over a given period. High service availability is vital for guaranteeing that users can rely on the provider for ongoing access to financial services.

Parameter 2: Performance Score

The Performance Score is a mix of both system measurements and user feedback. The system measurements include transaction latency, error rates, throughput, and other technical performance indicators, while user feedback gives subjective insights into how dependable the provider is regarded to be.

- System Metrics: These include transaction processing times, throughput (transactions per second), and error rates (% of unsuccessful transactions).
- User Feedback: Feedback from users on dependability, speed, and service quality is collected through questionnaires. The feedback is standardized to fit within the same range as the system measurements.

Parameter 3: Defects Score

This parameter tracks defects and also operational issues. System-generated alerts usually contain incidents of interest to key failure indicators on:

1. Compliance Breach: Indicates the violation of AML or KYC standards.
2. Fraudulent Activity: Flags for fraud attempts and suspicious behavior.
3. Transaction Failures: Bursts failed transactions due to system or human-generated errors.
4. System Downtime: It measures extended periods where the service is unavailable.
5. Delayed Transactions: Flags transactions that exceed acceptable processing time thresholds.

3.3. Layer III: Parameter Integration via FIS:

The FIS employs fuzzy logic rules to integrate parameters, transforming exact inputs into linguistic categories (e.g., Low, Medium, and High) and subsequently converting them into a definitive trust score. This architecture provides a thorough and clear assessment of trustworthiness.

Model Methodology

As well as the architecture of the model composed of three layers, the methodology of how the trust score is evaluated is also structured into three phases. The phases detail data collecting and preparation, trust parameters evaluation, and trust score calculation.

3.4. Phase 1: Data Collecting and Preparation

This phase aims to collect and prepare data from various sources so that it is valid and proper for analysis. where [22] emphasizes the importance of combining multiple data sources (as in Figure 1) to enhance trust assessments in complex environments through its contributions to development.

- Data Pre-processing and Normalization:
  Since metrics frequently have disparate units or ranges data normalization is crucial. Equation (1) is used to standardize all metrics and normalize data to a [0 1] scale.

$$X_{\text{normalized}} = \frac{X_{\text{raw}} - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

  Where $X_{\text{normalized}}$ is the metric value that has been normalized and scaled to [0, 1]. X_{raw} is the value of the raw data of the metric. $X_{\max}, X_{\min}$ are the dataset's lowest and highest recorded values for the metric.
  This step guarantees that every metric irrespective of its initial scale or unit of measurement makes an equal contribution to the ultimate score.
- Data Validation:
  To eliminate outliers or inaccurate data entries anomaly detection is essential. Due to its suitability for high-dimensional datasets such as financial transaction logs, the Isolation Forest algorithm is employed for this task. The Isolation Forest recursively partitions the data and chooses features at random to isolate anomalies. Where, for each data point, the algorithm computes an anomaly score using Equation (2):

$$A(x) = 2^{-\frac{E(h(x))}{c(n)}} \tag{2}$$

  Where: $A(x)$ is the anomaly score for data point x. $E(h(x))$ is the expected path length for x. $c(n)$ is the average path length of a binary tree with n observations. Anomaly scores closer to 1 indicate anomalies and closer to 0 indicate normal data points.
  This process helps identify data points that do not conform to the typical pattern thus preventing them from skewing the trust evaluation results.

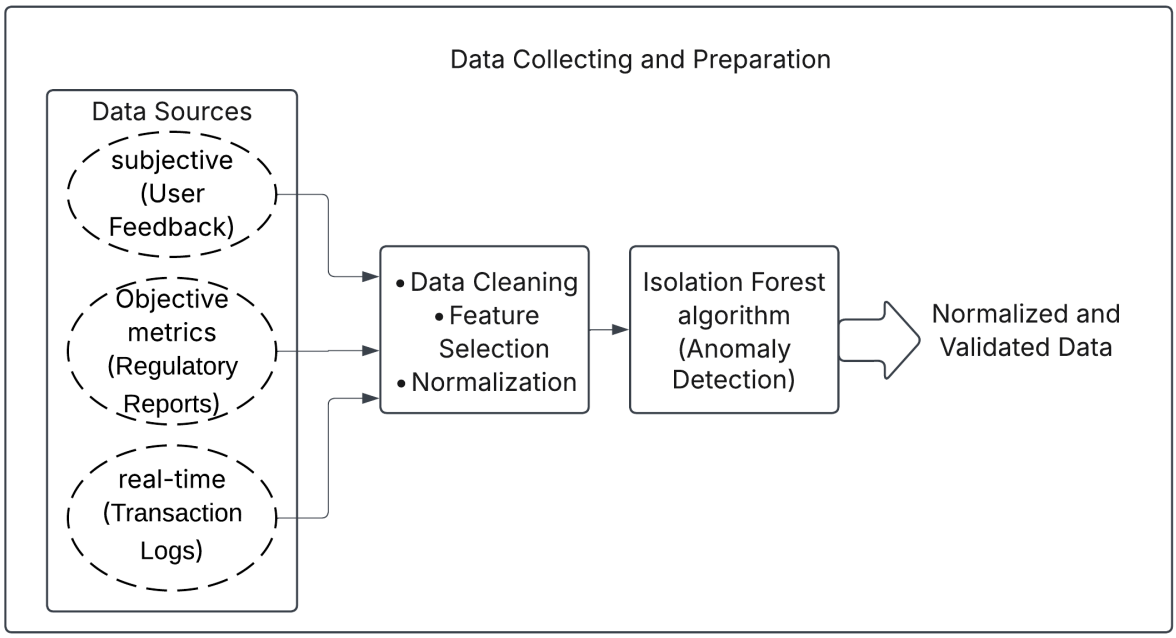

Figure 1. Data Sources and Preparation

3.5. Phase 2: Multi-Parameter Trust Evaluation

This phase examines three major trust parameters: FSLA Compliance, Performance, and Defects and Alerts Tracking. Each time one of these parameters is assessed adds to the ultimate trust score. The following algorithm describes how each parameter is evaluated.

Algorithm 1 Phase2_TrustParameterEvaluation

Input: NormalizedCleanedData
Output: FSLAComplianceScore, PerformanceScore, DefectScore
Step 1: Evaluate FSLA Compliance
$R_{\text{adherence}} \leftarrow$ Computed from regulatory data
$F_{\text{Prevention}} \leftarrow$ Analyze metrics
$T_{\text{Accuracy}} \leftarrow$ error rates from logs
$S_{\text{Availability}} \leftarrow$ uptime ratio
$S_{\text{FSLA}} \leftarrow \text{WeightedSum}(\text{R}_{\text{adherence}}, F_{\text{Prevention}}, T_{\text{Accuracy}}, S_{\text{Availability}})$
Step 2: Predict Performance Score
Combine $UserFeedback$ and system metrics (latency, throughput)
Train $GBMModel$ using labeled performance data
Apply grid search for model hyperparameter tuning
$S_{\text{Performance}} \leftarrow GBMModel.predict(NormalizedCleanedData)$
Step 3: Compute Defect Score
foreach *alert type* **do**
  $AlertScore_{\text{Weighted}} \leftarrow Weight \times AlertCount$
$D_{\text{Total}} \leftarrow \text{Sum}AlertScore_{\text{Weighted}}$ for all alert types)
return S_{FSLA}, $S_{\text{Performance}}$, D_{Total}

Parameter 1: FSLA Compliance Scoring

Each of the metrics (regulatory adherence, fraud prevention, transaction accuracy, and service availability) is given a weight based on how important it is in the financial world. Regulatory adherence for instance is regarded as the most important component since it guarantees credibility in the financial industry.

$$S_{\text{FSLA}} = w_1 \cdot R_{\text{adherence}} + w_2 \cdot F_{\text{prevention}} + w_3 \cdot T_{\text{accuracy}} + w_4 \cdot S_{\text{availability}} \tag{3}$$

Where (w_1, w_2, w_3, w_4) are the weights to each factor respectively, $R_{\text{adherence}}, F_{\text{prevention}}, T_{\text{accuracy}}, S_{\text{availability}}$ are the normalized values of the respective metrics, and S_{FSLA} is the FSLA Compliance Score.

A combination of proposed weights and an adaptive weighting method is used in our model to ensure both stability and flexibility in trust evaluation. Initially, the weights for each trust metric (e.g., FSLA compliance, performance, and defect tracking) are established based on expert knowledge or previous data. for example, regulatory adherence may start with a weight of 0.4, fraud prevention with 0.3, and so on. These weights give a baseline for the model. However, to accommodate for dynamic changes in regulatory and market situations, an adaptive weighing system is added. The weight of each parameter i at time t, abbreviated as $w_i(t)$, is updated using the following equation:

$$w_i(t) = w_i(t-1) + \alpha \cdot \Delta R_i(t) + \beta \cdot \Delta M_i(t) \tag{4}$$

Where:

- $w_i(t-1)$: The weight of parameter i at the previous time step.
- $\Delta R_i(t)$: The change in regulatory importance for parameter i at time t, derived from regulatory updates or compliance reports.
- $\Delta M_i(t)$: The change in market conditions affecting parameter i at time t, calculated from market volatility indices or transaction trends.
- α, β: Tuning parameters that control the impact of regulatory and market changes on the weights.

Parameter 2: Performance Evaluation

A Gradient Boosting Machine (GBM) is utilized to integrate the system metrics and user feedback and then forecast the Performance Score. GBM is chosen because it handles both categorical and continuous data adequately, making it suited for integrating system metrics and user input. Nonetheless, hyperparameter tuning is used to increase the GBM's predictive efficiency and accuracy. The model's primary parameters including the number of trees learning rate maximum depth and minimum child weight are optimized through hyperparameter tuning which guarantees that the model performs well when applied to new data.

In order to systematically explore parameter combinations and choose the configuration that minimizes the error metrics of the model such as mean squared error (MSE) the hyperparameter tuning process combines cross-validation with a grid search method. The model parameters are such as the maximum depth which controls the complexity of each decision tree and the learning rate which dictates how much the model modifies weights with each iteration. The Performance Score is computed as follows:

$$S_{\text{Perf}} = \text{GBM}_{\text{Model}}(U_{\text{feedback}}, S_{\text{metrics}}) \tag{5}$$

Where $\text{GBM}_{\text{Model}}$ is the trained model, U_{feedback} is the user feedback data, S_{metrics} represents the system metrics, and S_{Perf} is the performance score.

Parameter 3: Defects Tracking

Alerts such as compliance breach, fraudulent activity, transaction failure, system downtime, and delayed transactions are weighted to compute the Defect Score:

$$D_{\text{Total}} = w_1 \cdot A_{\text{compliance}} + w_2 \cdot A_{\text{fraud}} + w_3 \cdot A_{\text{failure}} + w_4 \cdot A_{\text{downtime}} + w_5 \cdot A_{\text{delay}} \tag{6}$$

Where $(w_1, w_2, w_3, w_4, w_5)$ are the weights that are assigned to alert types to reflect their relative impact on trust.

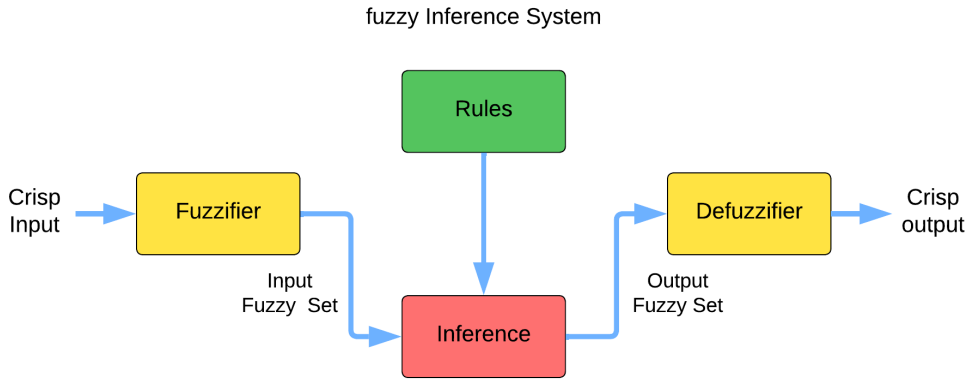

Figure 2. Fuzzy Inference System Procedure.

3.6. Phase 3: Comprehensive Trust Score Calculation

In this phase, FIS provides a systematic way to evaluate trustworthiness by examining the parameters calculated in the second phase and integrating them using preset fuzzy logic rules. The FIS runs on the concepts of fuzzy logic, which simulates the uncertainty and unpredictability inherent in subjective and objective judgments. By changing crisp input values (e.g., compliance percentages, performance metrics, and defect scores) into fuzzy sets, the FIS may reason via fuzzy rules to provide an output trust score in language words such as Low, Medium, High, or Excellent. The third stage, defuzzification, transforms these verbal outputs into a precise numerical trust score (the crisp output value) as shown in Figure 2.

FIS Inputs and Membership Functions:

The FIS processes three essential inputs (FSLA Compliance Score, Performance Score, and Defect Score) These inputs are standardized to a [0, 100] range to maintain consistency and compatibility with the fuzzy inference framework.

To model the trust evaluation process properly, the trapezoidal membership functions are applied to the inputs while triangular membership functions are applied to the output. The trapezoidal membership function $\mu(x)$ for FSLA inputs is represented in Equation (7):

$$\mu(x) = \begin{cases} 0, & \text{if } x \leq a \text{ or } x \geq d \\ \frac{x-a}{b-a}, & \text{if } a < x \leq b \\ 1, & \text{if } b < x \leq c \\ \frac{d-x}{d-c}, & \text{if } c < x \leq d \end{cases} \tag{7}$$

Where a, b, c, d are the parameters of the trapezoidal membership function.

Triangular functions divide the Trust Score into understandable terms (Low, Medium, High, and Excellent). In contrast, trapezoidal functions provide flexibility and smooth transitions when defining fuzzy sets (Low, Medium, and High). In addition to providing clear actionable trust evaluations in the output, these functions guarantee clear-cut modeling of input variability. This method preserves decision-making accuracy and interpretability while accounting for uncertainty in financial data.

Fuzzy Logic Rules:

The FIS utilizes a set of fuzzy logic rules developed based on expert knowledge and domain-specific insights to translate combinations of input memberships to output memberships. These rules capture the connections between the inputs and their effect on the trust score. FIS employs 27 rules to evaluate the trust score, for brevity, a subset of these rules is presented below:

- **IF** FSLA is **High** , Performance is **High**, and Defect is **Low** **THEN** Trust is **Excellent**.
- **IF** FSLA is **High**, Performance is **Medium**, and Defect Score is **Low** **THEN** Trust is **High**.
- **IF** FSLA is **Medium**, Performance is **High**, and Defect Score is **High** **THEN** Trust is **Medium**.
- **IF** FSLA is **Low**, Performance is **Medium**, and Defect Score is **High** **THEN** Trust is **Low**.
- **IF** FSLA is **Medium**, Performance is **Medium**, and Defect Score is **Medium** **THEN** Trust is **Medium**.
- **IF** FSLA is **High**, Performance is **Low**, and Defect Score is **Medium** **THEN** Trust is **Medium**.

These regulations signify the comparative significance of each parameter. For example, FSLA Compliance is prioritized due to its essential function in maintaining regulatory conformity, whereas Defect Score adversely affects the trust score, as operational problems such as fraud or downtime diminish reliability.

Defuzzification

Following the evaluation of the fuzzy rules, the FIS produces output membership values that fall into the Low, Medium, High, or Excellent trust score categories. Defuzzification is done using the centroid method, which yields a clear numerical trust score. The centroid method determines the fuzzy membership curve's center of the area under it.

$$T_{\text{Overall}} = \frac{\sum_{i=1}^{n} \mu_i \cdot x_i}{\sum_{i=1}^{n} \mu_i} \tag{8}$$

Where: μ_i is the value of membership, and x_i represents the crisp values. Both interpretability (as linguistic categories) and actionability (as a numerical trust score) are guaranteed by this process.

In the next section a representation of how the proposed methodology is implemented and the result of executing the implemented model.

4. Implementation and Results

In this section, we present the implementation procedure step-by-step of the proposed Trust Evaluation Model for Financial Providers using Fuzzy Inference Systems (FIS). Also, represents the results of executing the model.

Implementation

The proposed model of trust assessment was implemented in Python and its libraries for machine learning aspects, such as scikit-learn, and SciPy for the mathematical functions. FIS was designed with a mix of trapezoidal and triangular membership functions to represent uncertainty and linguistic variables. Basic rules governing the FIS were defined based on expert knowledge and fine-tuned with the help of simulation data.

We set up the Isolation Forest technique for handling high-dimensional data in anomaly identification by showing outliers both in the transaction logs and regulatory compliance reports. GBM utilizes performance evaluation, optimized using a grid search and cross-validation so that the efficacy of its prediction is ensured. Here is the implementation for each phase:

4.1. Phase 1:

Collecting the input data sources which are user feedback, transaction logs, and regulatory reports. These types of data are essential for evaluating the reliability of financial service providers since they record user's opinions, business operations, and compliance with regulations. In practical implementation, it is expected that these data

sources are accessible since financial institutions usually keep thorough records for compliance and operational reasons.

However, throughout the development stage, real-world data was inaccessible owing to privacy and accessibility concerns. Simulated models created to replicate reasonable financial situations helped to verify the architecture. Simulated data guaranteed the resilience of the model by allowing testing over a broad spectrum of situations. That data includes fictional but realistic values for user comments, transaction logs, and regulatory compliance measures. For instance:

1. User feedback ratings: These are provided on a Likert scale (1–5), normalized to a range of [0, 1].
2. Transaction logs: These include processing times or error rates following statistical distributions observed in similar systems.
3. Compliance data: These comprise scenarios with varied degrees of adherence, from flawless compliance to regular infractions.

After the dataset was generated and normalized, the Isolation Forest algorithm was applied to identify and remove outliers from the dataset.

4.2. phase 2

Before going to the trust parameter evaluation process, the adaptive weighing technique is implemented by continuously monitoring regulatory updates and market conditions. For each trust parameter, $\Delta R_i(t)$ is determined based on the frequency and severity of regulatory changes, whereas $\Delta M_i(t)$ is obtained from real-time market data, such as transaction volumes or system performance indicators. The weights are updated regularly to make sure the trust model stays up-to-date with the latest rules and market changes. For instance, if a new rule to stop fraud is set up, the weight for fraud prevention in the FSLA score will increase, showing it's now more important. This automatic change works together with the starting weights, which gives the model a stable foundation to build on. since we used simulated data, we used proposed weights in the computations as will be shown in this phase where the trust parameter evaluation process is implemented as follows:

FSLA Compliance Score: This score was determined using data values of fraud prevention, transaction accuracy, service availability, and regulatory compliance. Each element was standardized and given a weight according to its importance to the financial services sector; regulatory conformity received the highest weight 0.4. Then fraud prevention with a weight of 0.3, and transaction accuracy and service availability with weights of 0.2,0.1 respectively.

Performance Score: A Gradient Boosting Machine (GBM) was built, it combines system metrics and user feedback to predict the Performance Score. This is a thorough explanation of how the GBM was built.

- Data Preparation: data of the system metrics and user feedback was generated with features transaction latency, throughput, and error rates for the system metrics. And survey ratings that represent how users feel regarding the provider's service reliability.
- Hyperparameter Tuning: hyperparameters were tuned to optimize the model's performance. Learning rate, Number of trees, and Maximum depth of trees are the parameters that were tuned. Grid search and cross-validation were used for this process to decrease the mean squared error (MSE) of forecasts.
- Model Training: the simulated data for this part was divided into a training set (80% of the data) and a testing set (20% of the data). The training step helps the model to understand the data and features which help the model to predict a thorough Performance Score.

Defect Score: Since there is a set of alerts specified in the previous section, the simulated data was generated related to these alerts (compliance breaches, fraud detection, transaction failures, system downtime, and delayed transactions). Also, as FSLA compliance score a weighted sum function was used with the weights of (0.3, 0.3, 0.2, 0.1,0.1) respectively.

4.3. phase 3

In this phase, the FIS produces the overall trust value of a service provider by using the parameters (FSLA Compliance Score, Performance Score, and Defect Score) as inputs. During the fuzzification process, the values of these parameters are transformed into fuzzy linguistic input variables, creating the input fuzzy set (Low, Medium, High) as shown in Table 1. The trust level was assessed using the predefined fuzzy rules. And it is mapped to the combination of these inputs by the fuzzy rules. A clear Trust Score was computed on a scale of 0 to 100 by defuzzifying the results using the centroid method after applying the fuzzy rules.

Type of FIS	Mamdani	
Type of Defuzzification	Centroid	
Membership Functions	**Inputs**	**Output**
	- FSLA Compliance Score - Performance Score - Defects Score	- Trust Score
No. of Membership Functions	3	4
Type of Membership Functions	Trapezoidal	Triangle
Membership Function Names	Low, Medium, High	Low, Medium, High, Excellent
Range	0–40, 30–70, 60–100	0–30, 30–60, 60–90, 90–100

Table 1. Fuzzy Inference System Details

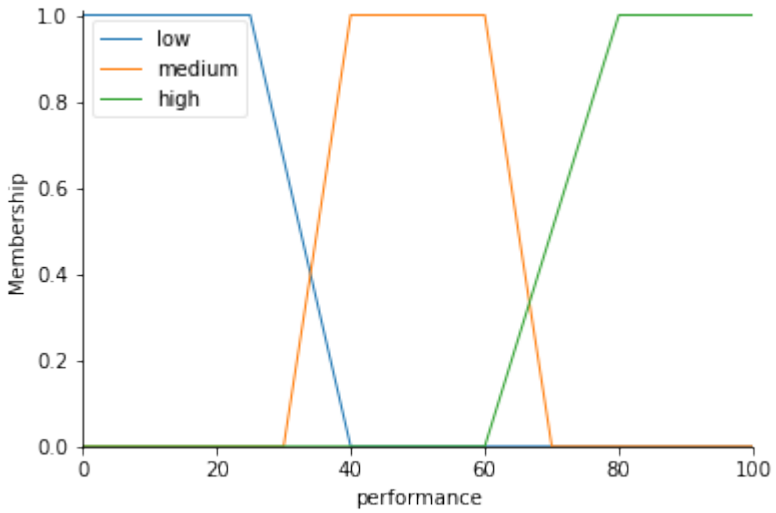

Figure 3. Input fuzzy sets.

The trapezoid-curve shape was utilized to represent the relative fuzzy membership functions of the three fuzzy inputs (Low, Medium, High) of the three input parameters as illustrated in Figure 3, As already mentioned, to determine the service provider's crisp overall trust value, centroid defuzzification was carried out. The output value membership functions (low, medium, high, excellent) are represented by the triangular-shaped curve as in Figure 4.

The overall trust value of the service provider falls into one of the four fuzzy output sets. To calculate the total trust level of a service provider, taking into account its FSLA Compliance Score, Performance Score, and Defect Score, the used rules were introduced to the Mamdani inference system. Each rule establishes a correlation between the input and output variables.

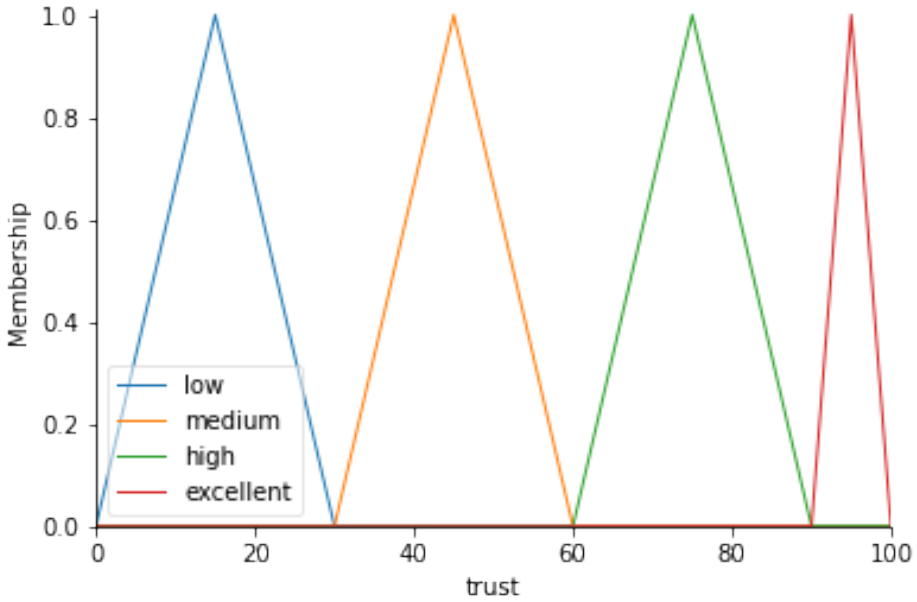

Figure 4. Output fuzzy sets.

Results

With the model implemented, five different datasets were generated to be used to test the model, the simulated data representing Five distinct financial service providers. The outcomes of computing each parameter in phase 2 are the input of the FIS which produces the overall trust value and status. Table 2 shows the values of each trust parameter and also the trust score and status resulting from executing the implemented model.

Provider	FSLA Compliance	Performance	Defects	Trust Score	Trust Status
Provider A	85	70	10	95.00	Excellent
Provider B	77	67.5	35	80.95	High
Provider C	56.8	35.7	53.9	52.02	Medium
Provider D	46	42	70	15.00	Low
Provider E	33.7	87.3	20.4	43.57	Medium

Table 2. Trust Parameters and Overall Trust Score for Providers

The results indicate that providers with varied degrees of performance, compliance, and flaws may be distinguished using the Trust Evaluation Model as in Figure 5. In order to test how well our model performs under different conditions, the simulated data represented four different scenarios. Where the data was created to highlight real-world problems that may face the financial service providers. The different scenarios are described in the model results which summarized in Table 2 are as follows:

- Normal Operations: This would be the situation where everything works as expected, meaning that regulatory compliance, performance of systems, and tracking defects are all within acceptable limits. The trust evaluation model should give the highest trust scores for providers operating under this condition as for Provider A.
- Compliance Breaches: This is a hypothetical situation that allows a provider not to meet such vital regulatory necessities as AML or KYC standards. It needs the model to perceive the lack of compliance and show the risk of this provider based on the scores of trust he receives as for provider B.

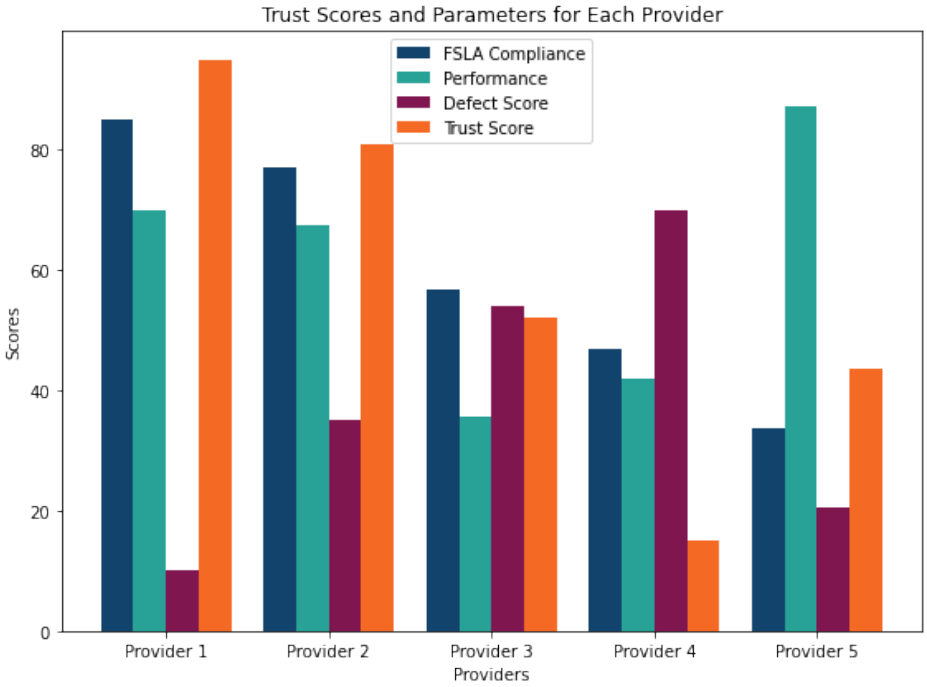

Figure 5. Model results and trust score and status of each provider.

- Fraudulent Activities: This would represent cases where irregularities or abnormalities in the information about transactions signal fraud—for example, unauthorized transactions or situations of security breach. In such cases, the model is put to task to identify and penalize such activities such as for provider D.
- Mixed Cases: A mix of compliance breaches and operational failures coupled with attempts to commit fraud. Most complex because the model needs to evaluate several variables at once and balance the trust score such as for providers C and E.

Testing across these four distinct scenarios helps determine if the model:

- Performs well in normal conditions (i.e., high trust scores for compliant providers).
- Accurately identifies and penalizes non-compliance or operational failures.
- Can adapt to and handle complex situations where multiple issues occur at once.

This would mean that the model will perform effectively in different scenarios of the real world by properly testing its correctness, flexibility, and reliability. Thus, a set of performance metrics were used to assess the model's effectiveness. The results of metrics (Accuracy, Precision, Recall, F1-Score, and ROC-AUC) are summarized in the table 3 :

Scenario	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
Normal Operations	95.2	93.8	96.5	95.1	98.3
Compliance Breaches	91.4	89.6	92.3	90.9	95.7
Fraudulent Activities	88.7	85.2	90.5	87.7	93.1
Mixed Cases	90.3	88.1	91.7	89.8	94.6

Table 3. Performance metrics across different scenarios

The results depicted that, in all tested scenarios, the model held an accuracy above 90%. Moreover, the ROC-AUC (Receiver Operating Characteristic—Area Under Curve) has always been higher than 90%, indicating a very strong discrimination capability to tell trustworthy providers from untrustworthy ones.

- Normal Operations: The highest performance was recorded under this scenario; the trust parameters were stable and passed all the thresholds that were expected, reflecting a well-functioning system.
- Compliance Breaches: The model turns out to be pretty effective in the detection of providers not complying with AML and KYC standards, yielding as high as 92.3% recall of all the violations.
- Fraudulent Activities: Though anomalies are artificially introduced into transaction logs, the model puts up a decent performance with an accuracy of 85.2%, hence reducing the chances of false positives.
- Mixed Cases: In complicated scenarios involving a combination of failures, the model obtained balanced metrics, showing its robustness and adaptability.

In real-world applications, financial institutions can use the system to acquire insights into which elements of their operations require improvement, whether it be strengthening compliance procedures, boosting system performance, or eliminating flaws. Additionally, this approach delivers an actionable decision-making tool for stakeholders, regulators, and customers by delivering clear and intelligible trust scores that reflect both the legal compliance and operational success of providers.

Comparative Analysis

The proposed model of trust evaluation extends the traditional frameworks in a number of key areas: adaptability, comprehensiveness, and reliability. Each of these enhancements is discussed below:

1. Real-Time Scoring
   This work integrates FIS with machine learning that could be used for dynamic trust evaluations by the model, considering it will have a real-time adaptation capability. This is one of the very important capabilities missing in most of the traditional systems that usually rely on static evaluations. For example, [7] indicated the importance of dynamic trust systems in fog computing environments, so this model aims to include real-time adaptability. Similarly, [5] showed the importance of dynamic updates in decentralized systems such as fog computing, further solidifying the necessity for real-time scoring in a trust evaluation framework.
2. Multi-Parameter Integration
   Unlike typical models that focus on a single facet of trust, the suggested framework integrates numerous aspects, including FSLA compliance, performance indicators, and defect tracking. This multi-parameter technique ensures that trust evaluations are comprehensive and granular. [22] highlighted the necessity of merging multiple data sources in trust evaluations to improve dependability, which is shown in the model's integration of operational, compliance, and user feedback metrics. Further, [11] has also proposed a hierarchical trust framework that has inspired the organization and weighting of these parameters within the model.
3. Fraud Detection
   The model employs the Isolation Forest method to identify fraudulent actions and anomalies in transaction data. This method lowers false positives quite successfully and improves the accuracy of trust assessments. By spotting abnormal patterns in sensor networks, which fits the application of Isolation Forest in this framework, the efficiency of anomaly detection techniques was proved by [20] in enhancing trust systems. Also, the fraud prevention measures within the integration align with evidence-based approaches by [12], which consider objective metrics blended with subjective insights for a robust fraud detection system.

These enhancements bring to light the ability of the proposed model to handle critical limitations in existing trust evaluation systems. The framework incorporates advanced techniques such as fuzzy logic, machine learning, and anomaly detection for a robust, interpretable, and adaptive solution in trust evaluation for financial service providers. These improvements enhance not only the reliability of trust assessments but also actionable insights for the stakeholders of the financial industry.

5. Conclusion

This work provides a holistic trust evaluation model for financial service providers, which holds significant potential for application in other domains such as e-commerce, healthcare, and the Internet of Things (IoT), that tackles important issues in the trust assessment for this domain. Using the combination of Fuzzy Inference Systems (FIS) and machine learning models, the proposed model aggregates FSLA compliance measures, operational performance information, as well as subjective user ratings to calculate dynamic, multi-dimensional trust scores. The model takes advantage of both the strengths of fuzzy logic in handling uncertainty and ambiguity, as well as the predictive power and real-time robustness of machine learning.

The findings demonstrate the effectiveness of this hybrid method in overcoming the constraints of existing trust evaluation frameworks, such as their static nature, reliance on subjective evaluations, and lack of integration across crucial variables. The model's capacity to react dynamically to changes in regulatory compliance, operational dependability, and fraud detection makes it a strong solution for the increasing needs of the financial industry. Moreover, the quantitative evaluation indicated good accuracy, precision, and recall, highlighting the model's reliability and practical application. These contributions underline the model's potential to promote transparency, accountability, and trustworthiness in financial services.

Future Work

Although the proposed model addresses the substantial gaps in current trust assessment frameworks, several aspects must be explored further in order to provide more scope and utility for it:

- Integration of External Data Sources: Future studies will focus on incorporating external data, such as market trends, social media sentiment analysis, and third-party compliance reports. These additional inputs can increase the model's comprehensiveness and offer a more holistic trust assessment.
- Real-World Deployment and Testing: Implementing the model in operational financial systems will evaluate its scalability and resilience under real-world situations. Field testing with multiple datasets will provide better insights into its efficacy across different financial environments.
- Advanced Machine Learning Techniques: Moving to advanced algorithms along the lines of deep learning or reinforcement learning would further improve this model's predictive accuracy with an ability to adapt itself in a complex, dynamic financial ecosystem. Also, it can be used to automatically tune the parameters and wehight α and β based on historical data and real-time feedback
- Scalability and Real-Time Performance: handles increasing data volumes and proposes potential optimization techniques, such as parallel processing and cloud-based architectures.
- User Interface and Real-Time Feedback: Designing a user-friendly interface would enable the practical application of the model by stakeholders. Additionally, implementing real-time feedback systems will enable continual improvement and refining of the trust evaluation process.
- Cross-Domain Applicability: Exploring the use of this model in other domains, such as e-commerce, healthcare, and IoT, could increase its value and demonstrate its adaptability to varied trust evaluation difficulties.

By addressing these areas, the suggested framework can evolve into a more versatile and powerful instrument for trust evaluation, thereby enhancing its contributions to the financial industry and beyond.

REFERENCES

1. R. Ali, M. Meraj, and M. S. Mubarik, *In the pursuit of financial innovation-led financial inclusion: A proposed construct for financial trust*, Borsa Istanbul Review, vol. 23, no. 6, pp. 1399–1413, 2023. https://doi.org/10.1016/j.bir.2023.09.002
2. C. Esposito, A. Galli, V. Moscato, and G. Sperlì, *Multi-criteria assessment of user trust in social reviewing systems with subjective logic fusion*, Information Fusion, vol. 77, pp. 1–18, 2021. https://doi.org/10.1016/j.inffus.2021.07.012
3. F. H. L. Chong, *Enhancing trust through digital Islamic finance and blockchain technology*, Qualitative Research in Financial Markets, vol. 13, no. 3, pp. 328–341, 2021. https://doi.org/10.1108/qrfm-05-2020-0076

4. M. B. Mansour, T. Abdelkader, M. Hashem, and E. M. El-Horbaty, *An integrated three-tier trust management framework in mobile edge computing using fuzzy logic*, PeerJ Computer Science, vol. 7, p. e700, 2021. https://doi.org/10.7717/peerj-cs.700
5. F. H. Rahman, T. Au, S. S. Newaz, W. S. Suhaili, and G. M. Lee, *Find my trustworthy fogs: A fuzzy-based trust evaluation framework*, Future Generation Computer Systems, vol. 109, pp. 562–572, 2018. https://doi.org/10.1016/j.future.2018.05.061
6. H. Khan, G. Chan, and F. Chua, *A fuzzy model for detecting and predicting cloud quality of service violation*, Journal of Engineering Science and Technology, vol. 13, pp. 58–77, 2018.
7. S. O. Ogundoyin and I. A. Kamil, *A Fuzzy-AHP based prioritization of trust criteria in fog computing services*, Applied Soft Computing, vol. 97, p. 106789, 2020. https://doi.org/10.1016/j.asoc.2020.106789
8. J. Luo, Z. Wei, J. Man, and S. Xu, *TRBoost: A generic gradient boosting machine based on trust-region method*, Applied Intelligence, vol. 53, no. 22, pp. 27876–27891, 2023. https://doi.org/10.1007/s10489-023-05000-w
9. A. Khalil, N. Mbarek, and O. Togni, *Fuzzy Logic Based Security Trust Evaluation for IoT Environments*, International Conference on Computer Systems and Applications (AICCSA), 2019. https://doi.org/10.1109/aiccsa47632.2019.9035294
10. F. Hendrikx, K. Bubendorfer, and R. Chard, *Reputation systems: A survey and taxonomy*, Journal of Parallel and Distributed Computing, vol. 75, pp. 184–197, 2014. https://doi.org/10.1016/j.jpdc.2014.08.004
11. S. Chong, J. Abawajy, I. Hamid, and M. Ahmad, *A multilevel trust management framework for service oriented environment*, In: International Conference on Innovation, Management and Technology Research, Malaysia, 2013, pp. 22–23.
12. A. Selvaraj and S. Sundararajan, *Evidence-based trust evaluation system for cloud services using fuzzy logic*, International Journal of Fuzzy Systems, pp. 1–9, 2017.
13. A. A. Battah, Y. Iraqi, and E. Damiani, *A Trust and Reputation System for IoT Service Interactions*, IEEE Transactions on Network and Service Management, vol. 19, no. 3, pp. 2987–3005, Sept. 2022. https://doi.org/10.1109/TNSM.2022.3179875
14. N. Fan, S. Shen, C. Wu, and J. Yao, *A hybrid trust model based on communication and social trust for vehicular social networks*, International Journal of Distributed Sensor Networks, vol. 18, 2022. https://doi.org/10.1177/15501329221097588
15. V. Cherkassky, *Fuzzy Inference Systems: A Critical Review*, Springer eBooks, pp. 177–197, 1998. https://doi.org/10.1007/978-3-642-58930-0_10
16. R. K. Chahal and S. Singh, *Fuzzy rule-based expert system for determining trustworthiness of cloud service providers*, International Journal of Fuzzy Systems, pp. 1–17, 2016.
17. V. Anandan and G. Uthra, *Defuzzification by area of region and decision making using Hurwicz criteria for fuzzy numbers*, Applied Mathematical Sciences, vol. 8, pp. 3145–3154, 2014. https://doi.org/10.12988/ams.2014.44294
18. J. Wang, X. Jing, Z. Yan, Y. Fu, W. Pedrycz, and L. T. Yang, *A survey on trust evaluation based on machine learning*, ACM Computing Surveys, vol. 53, no. 5, pp. 1–36, 2020. https://doi.org/10.1145/3408292
19. J. Liang, M. Zhang, and V. C. M. Leung, *A reliable trust computing mechanism based on multisource feedback and fog computing in social sensor cloud*, IEEE Internet of Things Journal, vol. 7, no. 6, pp. 5481–5490, 2020. https://doi.org/10.1109/JIOT.2020.2981005
20. R. Wu, X. Deng, R. Lu, and X. Shen, *Trust-based anomaly detection in emerging sensor networks*, International Journal of Distributed Sensor Networks, vol. 2015, pp. 1–14, 2015. https://doi.org/10.1155/2015/363569
21. C. Xin, X. Xianda, J. Yiheng, and W. Chen, *The trust evaluation and anomaly detection model of industrial control equipment based on multiservice and multi-attribute*, Proceedings of the International Conference on Computer and Communications, pp. 1575–1581, 2021. https://doi.org/10.1109/iccc54389.2021.9674285
22. E. Blasch, *Trust metrics in information fusion*, Proceedings of SPIE, 2014. https://doi.org/10.1117/12.2050255