An Enhanced Approach for Workmen's Compensation Insurance Fraud Detection Based on Fuzzy Rule-Based System

Reham M. Essa
Department of Information Systems,
Higher Institute for Management and Information Technology, Kafr El-Sheikh, Egypt

***Abstract*—In Workmen's Compensation insurance, fraud detection (FD) remains a significant challenge due to claims' inherent uncertainty and complexity. To address this, we propose an enhanced approach based on a fuzzy rule system (FRS) for FD. The FRS is designed to handle ambiguous and imprecise data, making it effective for identifying fraudulent patterns in insurance claims. Unlike traditional methods, the fuzzy system utilizes human-like reasoning by applying flexible rules to assess the likelihood of fraud under uncertain conditions. By modeling the decision-making process with fuzzy logic, the system allows for a detailed evaluation of claims, accommodating the gray areas that often exist in FD. This approach enables accurate and adaptive FD, reducing false positives and enhancing the precision of fraud identification. In imbalanced FD scenarios, the system achieves strong performance, such as an F1-score of 0.82 and MCC of 0.75, demonstrating its capability to correctly identify rare fraudulent cases despite class imbalance.**

***Keywords*—Workmen's compensation; fuzzy logic; fraud detection; rule-based system; insurance fraud; prediction**

I. Introduction

The insurance industry plays a vital role in supporting the economic stability of a country, mainly through employee compensation insurance, which provides coverage for workers injured on the job. This type of insurance helps cover wage loss, medical expenses, and other benefits related to work-related injuries. Nevertheless, it also exposes companies to potential fraud, as unethical employers, workers, and medical providers may exploit the system for illegal financial gain. Fraudulent activities can involve deliberate deception, misrepresentation, or violating laws and policies to obtain unauthorized benefits. Combating fraud requires sophisticated detection methods, as fraud can be challenging to identify due to its intentional manipulation of the system and the complex, often hidden, nature of these activities [1].

Fraud in Workmen's Compensation insurance can take many forms, each adding significant financial strain to the system and impacting both employers and employees. One common type is employee fraud, where individuals claim non-work-related injuries as work-related to receive benefits. This may involve exaggerating the severity of an injury or misrepresenting the timing of the injury. Employer fraud occurs when companies misclassify employees or underreport their payroll to lower insurance premiums, resulting in reduced contributions and increased risks for insurers. Healthcare provider fraud involves billing for unnecessary or nonexistent medical treatments, inflating costs, and further burdening the insurance system. Lastly, insurance carrier fraud can include wrongful denial of claims or manipulating information to avoid paying out legitimate claims. These types of fraud collectively undermine trust in the system, increase premiums for honest participants, and require sophisticated detection mechanisms to manage and mitigate their impact [2],[3].

A FRS is effective for handling uncertainty in detecting fraud in Workmen's Compensation insurance by allowing for flexible decision-making in cases where data is ambiguous or incomplete. Unlike traditional techniques that use rigid, binary classifications (fraudulent or not), fuzzy systems handle data in degrees, assigning partial membership to different categories, such as "somewhat suspicious" or "likely fraudulent." This flexibility enables the system to model real-world uncertainty accurately, reducing false positives and adapting more effectively to complex, evolving fraud patterns. Additionally, fuzzy systems incorporate expert knowledge and subjective judgment, translating it into actionable rules, which traditional methods struggle to do. This results in a more nuanced, precise, and adaptive FD system that is better suited for the multifaceted nature of insurance fraud [4]-[6].

Detecting fraud in Workmen's Compensation insurance is a challenging task due to the varying and complex nature of fraudulent activities. Fraud indicators, such as suspicious reporting times, data behavior patterns, case correlations, and demographic factors, are critical for insurance experts to detect fraud. On the other hand, subjective judgments from experts can lead to inconsistencies in decision-making, and managing large volumes of data manually is time-consuming, prone to errors, and ineffective in fraud prevention. Furthermore, the growing data collected by insurance companies is often underutilized, missing opportunities to detect hidden fraud patterns. To address the challenges of detecting fraud in Workmen's Compensation insurance, this research proposes a FRS for FD. Unlike traditional systems, a FRS can handle uncertainty and imprecision in data, making it especially useful for complex FD scenarios where decisions must be made based on incomplete or vague information. Fraud indicators like suspicious behavior, reporting time discrepancies, and demographic factors often involve degrees of uncertainty that rigid systems struggle to process.

Fuzzy logic allows the system to mimic human reasoning by evaluating these factors in a flexible manner, providing more nuanced and accurate FD. This approach improves decision-making by accounting for varying fraud patterns and expert judgments, minimizing the risk of errors and inconsistencies. Additionally, the system enhances the ability of insurers to process large amounts of data efficiently, reducing costs and extracting hidden patterns that manual analysis would likely miss. Ultimately, it supports long-term business sustainability by providing more reliable FD and better resource allocation for claim investigations [7],[8].

A. Problem Statement and Motivation

The primary problem in Workmen's Compensation insurance FD is the inherent uncertainty and complexity of claims, which makes it difficult to accurately identify fraudulent activities. Traditional FD methods fail to handle ambiguous data and often rely on rigid, binary decision-making processes, leading to either missed fraud cases or false positives. Fraud in this domain frequently involves subtle patterns and subjective factors, such as claim timing, injury severity, or inconsistent reporting, which can vary greatly between cases. To address these challenges and motivated by the need for more precise and adaptable detection systems, this research proposes a fuzzy rule-based approach that leverages the flexibility of fuzzy logic to handle imprecise data.

B. Novelty and Contribution

The novelty of the proposed research lies in its exclusive use of an FRS for detecting fraud in Workmen's Compensation insurance, which distinguishes it from traditional methods. While conventional FD techniques often rely on strict, binary classifications that struggle with ambiguous data, this research leverages fuzzy logic to handle the uncertainty and complexity inherent in insurance claims. The fuzzy system's ability to model human-like reasoning by applying flexible, adaptable rules is a significant advancement in the domain of FD. It addresses the common challenge of managing imprecise and subjective information, such as variations in injury reporting or inconsistent claim histories, which are often difficult to capture using rigid systems.

The main contribution of this research is the development of an intelligent, adaptive framework that reduces false positives and improves the precision of FD. By offering a more nuanced evaluation of claims, this fuzzy rule-based approach enhances decision-making capabilities and allows insurance companies to more effectively identify fraudulent patterns. This ultimately contributes to the sustainability and integrity of the Workmen's Compensation insurance system, ensuring fair claims processing and reducing unnecessary costs caused by fraud. The proposed system's ability to extract hidden patterns from data and adapt to changing fraud schemes is another critical advancement, positioning it as a valuable tool in minimizing fraud risk.

The rest of the study is organized as follows, in section II state of the arts. In section III, we discussed and presented the research methodology in detail. Section IV presented the experiment and a discussion of the results followed by a conclusion.

II. Related Works

In the field of Workmen's Compensation insurance FD, several state-of-the-art techniques have emerged, each with its advantages and disadvantages. These methods can be broadly classified into data mining techniques, machine learning algorithms, rule-based systems, and hybrid approaches that combine multiple models for enhanced FD [9]-[15]. Data mining techniques like clustering (e.g., K-Means) and association rule mining are widely used to group similar insurance claims and identify hidden patterns in large datasets. These methods can efficiently process vast amounts of data, helping detect fraud by flagging anomalies. However, they rely heavily on historical data, making them less adaptable to evolving fraud tactics, and they can lead to high false-positive rates, incorrectly identifying legitimate claims as fraudulent [9], [12],[13].

Machine learning algorithms, including supervised and unsupervised models like random forest, support vector machines (SVM), and Neural Networks, are increasingly popular for FD. These models excel at learning from new data, detecting complex and non-linear relationships between variables, and are effective at identifying fraud patterns with high accuracy when trained on quality datasets. However, they require large, well-labeled datasets for effective training, which can be difficult to acquire. Additionally, many machine learning models, particularly deep learning networks, function as black-box systems, making their decision-making process difficult to interpret, which can lead to trust issues in critical applications such as FD [16]-[19], [20]-[23].

Traditional rule-based systems use predefined rules, usually set by experts, to classify claims as fraudulent or non-fraudulent. These systems are transparent and easy to understand, as they follow clear, interpretable rules, allowing for quick decision-making and reliability in detecting known fraud patterns. However, they lack flexibility since they can only detect fraud based on the programmed rules, making them ineffective against new or evolving fraud schemes. This limitation often leads to false negatives, where fraudulent cases are missed because they don't fit the established patterns. FRS, an advanced version of traditional rule-based approaches, utilize fuzzy logic to manage uncertainty and imprecision in FD. They excel in cases where data is ambiguous or conflicting, allowing them to model complex human reasoning and effectively handle gray areas. This flexibility reduces false positives and improves the detection of subtle fraud patterns that rigid systems might overlook. However, designing and maintaining fuzzy systems can be more challenging, as their performance heavily relies on the quality of the crafted rules and the expertise required, which can introduce subjectivity [22]-[25].

Hybrid systems that combine machine learning models with rule-based systems or fuzzy logic offer a powerful approach to leveraging the strengths of each method. By integrating machine learning, these systems can adapt to emerging fraud patterns through data-driven insights and continuous learning, making them highly flexible and responsive. Simultaneously, the rule-based components bring transparency and domain expertise that make decision-making more interpretable and understandable, particularly in industries requiring accountability, such as

finance or healthcare. This balance between adaptability and transparency provides an advantage in terms of flexibility, accuracy, and control. However, these systems also come with significant challenges. Their complexity makes implementation difficult, often requiring extensive computational resources to manage the interactions between the machine learning and rule-based components effectively. Additionally, maintaining a system that remains both transparent and accurate over time necessitates ongoing tuning and optimization, as misalignment between the components could compromise either its clarity or predictive power [12],[13], [20]-[23].

Pallavi A. et al. [26] carried out a detailed study that aimed to design an analytical framework for identifying potentially fraudulent health insurance claims. The methodology combined machine learning-based predictive modeling with a rule-based scoring system. Historical claims data were analyzed using data mining techniques to uncover patterns and correlations associated with fraud, enabling the model to assign probabilistic fraud scores to new claims based on similarities with known fraudulent behaviors. Alongside this, a set of predefined rules—such as abnormal treatment costs, high claim frequency, or unusual claimant behavior—was used to assign additional risk scores. Claims with higher cumulative scores were flagged for further investigation. This hybrid approach offered improved accuracy, scalability for large datasets, and adaptability through continuous learning. However, it also presented challenges, including dependence on high-quality data, limited flexibility in detecting new fraud patterns due to rigid rules, and the possibility of false positives, which could impact genuine claimants.

Kirlidog and Asuk [27] conducted a study on health insurance FD by focusing on anomaly classification techniques to distinguish fraudulent claims from legitimate ones. Their methodology centered on analyzing historical insurance data to identify patterns of irregularities that deviate from typical claim behavior. By applying data mining and statistical analysis methods, they were able to classify anomalies that often signify fraudulent activity, such as inconsistencies in billing, unusual claim frequencies, or improbable medical procedures. This approach allowed for the detection of fraud by recognizing outliers in the dataset rather than relying solely on predefined rules. The key advantages of this method include its ability to uncover previously unknown fraud patterns and its adaptability to different datasets. However, it also has limitations, such as a high dependency on the quality and completeness of the input data, and a potential for false positives where legitimate but atypical claims may be flagged as fraudulent, necessitating further manual review.

This study, presented in Ref. [24], introduced a novel hybrid methodology for detecting fraud in automobile insurance claims by integrating a Genetic Algorithm (GA)-enhanced Fuzzy C-Means (FCM) clustering technique with multiple supervised classification models. The approach begins by dividing the original insurance dataset into a test set and a training set. The training data undergoes FCM clustering, optimized using GA, to perform undersampling and create meaningful, balanced clusters. Test instances are then categorized into three classes: genuine, malicious, or suspicious, based on their proximity to these clusters. Genuine and clearly fraudulent claims are filtered out, while suspicious cases undergo further evaluation using four supervised classifiers: Decision Tree (DT), (SVM), Group Method of Data Handling (GMDH), and Multi-Layer Perceptron (MLP). Model training and validation are performed using 10-fold cross-validation to ensure reliability. The strengths of this method lie in its ability to handle imbalanced data through intelligent clustering, reduce false positives by isolating uncertain cases, and improve accuracy through ensemble analysis. However, the system's complexity and computational cost are potential drawbacks, and its performance is highly dependent on proper tuning of clustering parameters and classifier settings.

In Ref. [25], the authors suggested a machine learning-based approach for detecting insurance fraud using a supervised classification algorithm trained on historical claim data, which includes both legitimate and fraudulent cases. The model utilizes various features to differentiate between genuine and suspicious claims. Evaluated on real-world insurance data, the approach demonstrated high accuracy and was effective in identifying fraudulent claims that traditional rule-based methods often miss. Its main strengths include improved detection rates, automation of the fraud identification process, and reduced workload for human investigators. However, the approach relies heavily on the quality and representativeness of the training data, and its performance may decline if fraudulent behavior patterns change or if unseen types of fraud are introduced.

Workmen's FD is inherently complex due to the multifaceted and often hidden nature of fraudulent behavior. This necessitates the use of advanced data mining techniques capable of transforming fragmented and seemingly unrelated data into coherent, actionable insights. A foundational step involves integrating data from diverse sources into a centralized, structured data repository to facilitate comprehensive analysis. In this context, Helmut F., et al. [28] used an explainable attention-based neural network for claim FD. This model uses attention mechanisms within deep learning architectures to focus on the most relevant parts of the data, enhancing interpretability while maintaining high performance. The key advantage of this method is its transparency—decision-making can be better understood by auditors and investigators, fostering trust in AI-driven processes. However, it can be computationally intensive and requires substantial training data to perform effectively, particularly in complex, high-dimensional FD environments.

Additionally, Sadgali I. et al. [29] investigated the performance of various machine learning techniques in detecting financial fraud, comparing algorithms such as Decision Trees, Random Forests, SVM, and Neural Networks. The study highlighted that ensemble methods like Random Forests generally achieved better accuracy and robustness due to their ability to reduce overfitting and model nonlinear relationships. However, such models can lack interpretability, making it difficult to understand how fraud decisions are made. Similarly, Ryman, T. et al. [30] conducted a survey and benchmarking study on how AI and machine learning are transforming payment card FD. The research emphasized the growing use of real-time anomaly detection, deep learning, and hybrid AI systems in the financial industry. While these models significantly enhance detection capabilities and reduce false

positives, their implementation involves high computational costs, complex deployment requirements, and the need for continuous updating to keep pace with evolving fraud tactics.

A. Research Gap

Based on the comprehensive survey of current methodologies in Workmen's Compensation insurance FD, a notable research gap emerges in the underutilization of uncertainty-based fuzzy rule mining techniques—particularly in environments where data is ambiguous, imprecise, or incomplete. While fuzzy logic has been applied in some advanced rule-based systems to manage uncertainty, most existing implementations either rely on traditional, rigid rule-based approaches or focus heavily on machine learning and data-driven models that often operate as black boxes. These methods may lack the ability to effectively handle borderline cases or precise fraud patterns that do not conform strictly to binary classifications. This is a significant limitation in real-world fraud scenarios, where subtle cues, partial patterns, or conflicting indicators may be present. Incorporating fuzzy rule-based mining could allow for better modeling of these "gray areas" in claim behavior, facilitating more human-like reasoning and enhancing both accuracy and interpretability. Moreover, while hybrid models combining machine learning and rule-based systems have shown promising results in terms of accuracy and flexibility, they often neglect the integration of fuzzy logic into the rule-generation or decision-making process. This represents an untapped opportunity to bridge the gap between high-performance prediction and explainable decision-making. By embedding fuzzy rule mining into hybrid frameworks, researchers could develop systems that are not only adaptive and robust to emerging fraud tactics but also transparent and capable of providing interpretable justifications for their outputs.

III. Methodology

In the landscape of Workmen's Compensation insurance, detecting fraudulent claims is a complex task due to the high degree of uncertainty, incomplete information, and subjective evaluations often associated with the claims process. Traditional FD methods, which are mostly rule-based or statistical, frequently struggle to manage the gray areas inherent in such assessments, leading to high false positive rates or missed fraud cases. To address these challenges, we propose a comprehensive FD model centered on an FRS. The use of fuzzy logic allows for handling ambiguity and imprecision in claim-related data, mimicking human reasoning through flexible if-then rules. To further strengthen the model, we incorporate clustering techniques to discover hidden patterns and structure within the claims data before applying fuzzy classification.

This layered approach enhances the model's ability to adapt to complex fraud behaviors, offering a more nuanced, intelligent, and sustainable solution to FD in Workmen's Compensation insurance. The proposed model consists of three main stages: Preprocessing, Clustering, and Fuzzy Rule-Based Classification. Each stage is critical to ensuring that the system is accurate, adaptive, and capable of maintaining the integrity of the claims process. The following subsection outlines the purpose of each step in the proposed model and explains its specific function within the FD process. Each stage is described with implementation details to illustrate how it contributes to the overall effectiveness of the system. Together, these steps form a cohesive framework for accurately identifying fraudulent insurance claims. Fig. 1 illustrates the structured workflow of the proposed fuzzy rule-based classification system, detailing each stage from data preprocessing to the final classification of claims. The process encompasses feature selection, clustering, fuzzy rule definition, and inference application, ensuring a comprehensive and interpretable classification framework.

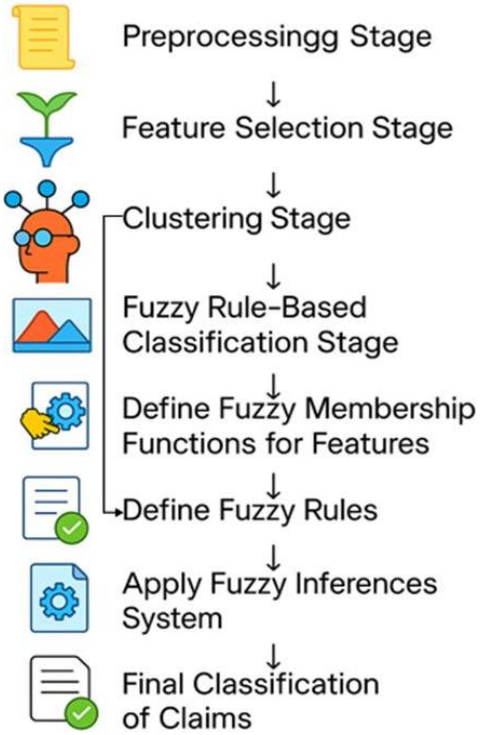

Fig. 1. Workflow of the proposed fuzzy rule-based classification system for claim processing.

A. Preprocessing Stage

The preprocessing phase is a critical foundation for building an effective FD model, as it transforms raw insurance claim data into a structured and analyzable format [24]. The first step, data cleaning, addresses inconsistencies such as missing or incomplete values—these can be handled through imputation techniques (in our case, K-nearest neighbors for numerical data, and the mode for categorical variables are employed). Duplicate records are also identified and removed to prevent redundancy and bias. Feature selection and extraction involve isolating the most informative attributes, such as claimant age, injury type, medical expenses, length of claim processing, and employer reputation. These features are often selected based on domain knowledge, correlation analysis, or dimensionality reduction methods like PCA (PCA is employed in the implementation). New features can also be engineered to provide more context—for example, calculating "reporting delay" (the time between incident and claim filing) or "claim frequency" (number of claims filed by the same individual or employer over a period) [25],[26].

In our case, the "insurance claims fraud data" dataset on Kaggle https://www.kaggle.com/datasets/mastmustu/insurance-

claims-fraud-data) is used. This benchmark dataset comprises 1,000 records and 40 features, offering a comprehensive view of various aspects related to insurance claims. Key features include 'policy_annual_premium', 'incident_type', 'collision_type', 'incident_severity', 'insured_hobbies', 'auto_make', 'auto_model', and 'auto_year', among others. The dataset also contains a target variable indicating whether a claim is fraudulent, facilitating supervised learning approaches. The diverse range of features allows for the exploration of complex relationships and patterns associated with fraudulent activities, making it a valuable resource for developing and testing FD models.

B. Feature Selection Stage

For implementing a fuzzy-rule-based classification in the Insurance Claims Fraud Dataset, the four most prevalent features suitable for capturing fraudulent activities would be total_claim_amount, incident_severity, property_damage, and fraud_reported. The selection of these features is based on expert knowledge within the insurance domain. Total_claim_amount is a critical feature because, in the context of FD, inflated or exaggerated claims are common indicators of fraudulent activity. Experts recognize that claims involving unusually high amounts, especially when compared to similar incidents, are more likely to be fraudulent.

Similarly, incident_severity is a key feature because the severity of the incident can provide significant insights into the fraud risk. Serious incidents often result in larger claims, but when paired with other anomalies like inflated damage reports, they can indicate potential fraud. Property_damage is another essential feature, as fraudulent claims may involve exaggerating or fabricating the extent of property damage. Insurance experts are well aware that discrepancies between reported damage and actual damage can be a red flag for fraudulent claims. Finally, fraud_reported, the target variable, directly correlates to the task of FD, serving as the benchmark for building fuzzy rules [18],[25].

These features were selected because they encapsulate the most common and easily identifiable signs of fraud, making them ideal for developing a robust fuzzy-rule FD system to classify and analyze claims with varying degrees of uncertainty, which is key to detecting nuanced fraudulent patterns. Once relevant features are selected, normalization or standardization is applied to numerical values to ensure consistency across varying scales, which is particularly important when combining data into a fuzzy system. Min-Max scaling Technique is employed for this purpose.

C. Clustering Stage

Clustering serves as a critical step in the proposed Workmen's Compensation insurance FD framework by identifying underlying patterns and structural relationships within claim data. Among available methods, C-Means clustering is particularly well-suited for this domain because it allows each claim to have varying degrees of association with multiple clusters. This soft assignment capability is essential in Workmen's Compensation cases, where the distinction between fraudulent and genuine claims is often blurred. C-Means effectively captures such ambiguity, enabling the model to handle borderline or suspicious cases more realistically.

C-Means algorithm offers several advantages when applied to FD in Workmen's Compensation insurance. First, it provides a straightforward way to categorize claims into distinct groups based on patterns such as claim amounts, recovery durations, or frequency of claims. By clustering similar claims together, the algorithm helps identify unusual patterns, such as clusters of claims with unusually high medical expenses, extended recovery times, or repeat claims by the same individuals. These anomalies can be indicative of potential fraudulent behavior, which makes C-Means valuable for initial FD in large datasets. Since each claim is assigned to a single cluster, the results are easy to interpret, providing a clear classification of claims into risk categories like low, medium, or high risk. Moreover, C-Means is computationally efficient and relatively simple to implement, making it suitable for real-time applications where speed is crucial. For Workmen's Compensation FD, the algorithm can effectively segregate claims into well-defined clusters, assisting fraud investigators in identifying and investigating suspicious patterns without the need for complex calculations or probabilistic models.

C-Means algorithm begins by selecting an optimal number of clusters (C), which is usually based on expert judgment (in our case low, medium, and high-risk categories). Initially, C random cluster centers are chosen from the data points. Each claim is then assigned to one of the clusters based on its proximity to the cluster centers. This is done by calculating the Euclidean distance between each claim and the cluster centers and assigning the claim to the closest cluster. After the initial assignment, the algorithm recalculates the cluster centers by taking the mean of all the claims assigned to each cluster. This process is repeated iteratively, with each iteration updating the cluster centers and reassigning claims to the closest cluster. The iterations continue until the cluster centers stabilize, meaning the changes in the cluster centers are minimal, or until a predefined number of iterations is reached. The final clusters can then be analyzed to identify patterns of fraud, such as outliers with excessive claims or suspiciously high medical costs, which might require further investigation. See [7],[24] for more details.

D. Fuzzy Rule-Based Classification Stage

In Workmen's Compensation FD, the output of the C-Means clustering stage can serve as a foundational input to the Fuzzy Rule-Based Classification stage by providing a clear categorization of claims into distinct clusters based on patterns of behavior. Once the claims are clustered (low, medium, and high-risk), these clusters can be mapped to fuzzy rules that define the characteristics of potentially fraudulent claims. For example, a cluster of claims with unusually high medical costs and extended recovery times can be labeled as "high-risk" and associated with fuzzy rules such as "IF medical costs are high AND recovery time is long, THEN the claim is likely to be fraudulent". The FRS then applies these rules to individual claims, using the degree of membership in each cluster (from the C-Means output) to assign a fuzzy value to the claim's likelihood of being fraudulent. The use of fuzzy logic allows for precise decision-making, where claims that don't strictly fit into one category (e.g., a claim that's borderline between high and medium risk) can still be assessed with a degree of uncertainty, making the FD system more flexible and accurate in identifying

suspicious claims [31]-[33]. Below is a detailed step-by-step breakdown of the Fuzzy Rule-Based Classification process.

1) Define fuzzy membership functions for features:

- Input: Data for each claim (e.g., total claim amount, incident severity, property damage, and fraud reported).
- For each of the selected features, fuzzy membership functions will be defined to describe the degree to which a claim belongs to categories. These functions are based on expert domain knowledge, which helps identify patterns indicative of fraud.
- The fuzzy membership function for the claim amount can be defined into categories such as "Low," "Medium," and "High." As shown in Fig. 2, this membership function reflects the common fraud pattern where unusually high claims are more likely to be fraudulent.

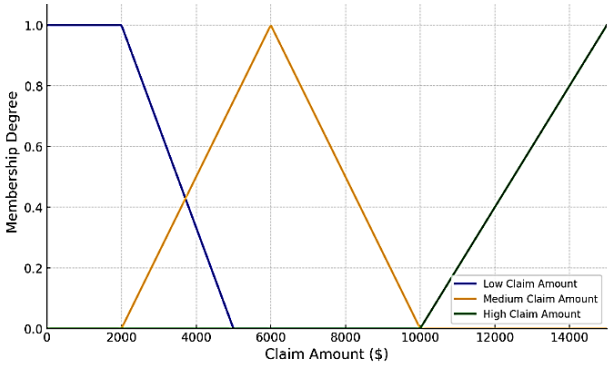

Fig. 2. Fuzzy membership functions for total claim amount.

- For incident severity, fuzzy categories like "Minor," "Moderate," and "Severe" are defined (see Fig. 3). This function helps capture the relationship between the severity of the incident and the likelihood of fraud, as more severe incidents often involve larger claims, but when coupled with other anomalies (e.g., inflated damage), they increase fraud risk.

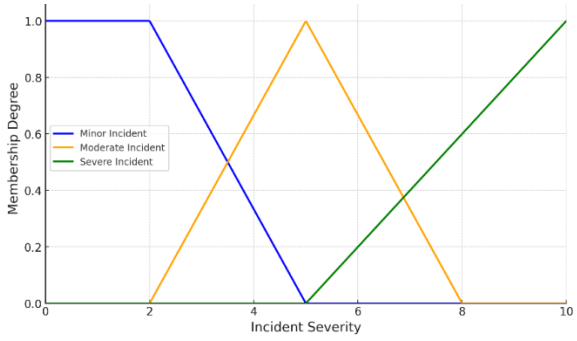

Fig. 3. Fuzzy membership functions for incident severity.

- The property damage feature could be fuzzified into categories such as "None," "Minor," and "Major." As shown in Fig. 4. Discrepancies between reported and actual property damage are common in fraudulent claims, and this membership function helps to capture such discrepancies. The fraud_reported feature is the target variable and would typically be fuzzified into two categories, "Fraud" and "No Fraud." (See Fig. 5). This fuzzy set provides the basis for determining the probability of fraud in each claim.

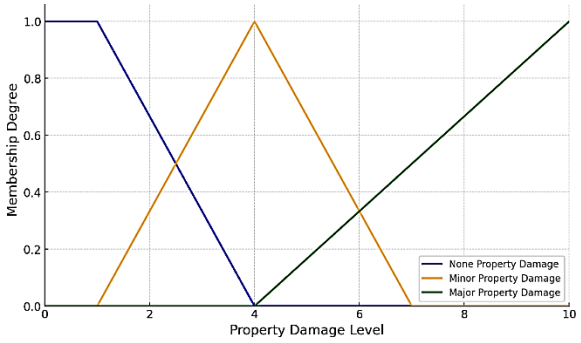

Fig. 4. Fuzzy membership functions for property damage.

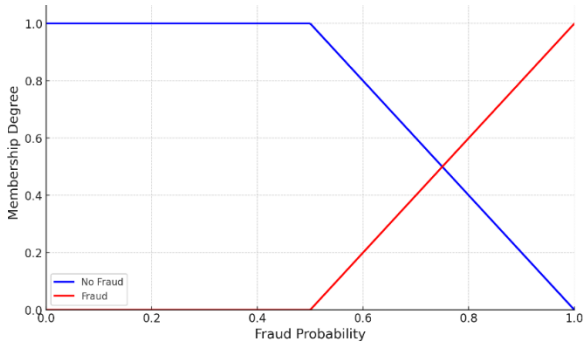

Fig. 5. Fuzzy membership functions for fraud reported.

- Output: The output consists of a set of fuzzy membership values for each claim, reflecting the degree that the claim belongs to the defined categories. For instance, a claim might have a membership value of 0.8 in the "High" category for claim amount, 0.2 in "Severe" for incident severity, and 0.5 in "Major" for property damage. These fuzzy values are used as inputs for the fuzzy inference system, which will evaluate and classify the fraud risk based on the fuzzy rules.

2) Define fuzzy rules:

- Input: The fuzzy membership functions created in the previous step, along with the C-Means clustering information.
- Based on expert knowledge or historical data, define fuzzy rules that link the fuzzy features to the risk of fraud. These rules describe the relationships between the membership values of various features and the likelihood of fraud.

Rule 1:

- IF (Claim Amount IS High) AND (Recovery Time IS Long) THEN (Fraud Risk IS High).
- Rule 1 reflects the assumption that high claim amounts and long recovery times are strongly associated with higher fraud risk. This is a typical fraud pattern, where

more severe incidents (e.g., costly claims and longer recovery periods) might be manipulated to gain higher payouts.

Rule 2:

- IF (Claim Amount IS Low) AND (Recovery Time IS Short) THEN (Fraud Risk IS Low).
- Rule 2 assumes that low claim amounts and short recovery times are indicative of legitimate claims with low fraud risk, as these types of claims tend to be more straightforward and less prone to exaggeration or manipulation.

Rule 3:

- IF (Claim Amount IS Medium) AND (Recovery Time IS Short) THEN (Fraud Risk IS Medium).
- Rule 3 reflects the case where medium claim amounts and short recovery times are still somewhat low in terms of fraud risk, but not as low as in Rule 2, hence the medium fraud risk.

Rule 4:

- IF (Claim Amount IS Low) AND (Recovery Time IS Medium) THEN (Fraud Risk IS Medium).
- Rule 4 assumes that low claim amounts with medium recovery times still do not strongly suggest fraud, but they are slightly more suspicious than very low amounts with short recovery times, leading to a medium fraud risk.

Rule 5:

- IF (Claim Amount IS High) AND (Recovery Time IS Medium) THEN (Fraud Risk IS High).
- Rule 5 shows that high claim amounts with medium recovery times are also indicative of high fraud risk, as the combination of a higher payout and moderate recovery times is a typical scenario for fraud.

Rule 6:

- IF (Claim Amount IS Medium) AND (Recovery Time IS Long) THEN (Fraud Risk IS High).
- Rule 6 suggests that medium claim amounts with long recovery times may also represent a high fraud risk, as long recovery times are more likely to coincide with inflated or extended claims to extract more money from the insurance.

Rule 7:

- IF (Claim Amount IS Medium) AND (Recovery Time IS Medium) THEN (Fraud Risk IS Medium).
- Rule 7 reflects that medium claim amounts with medium recovery times suggest a medium fraud risk due to the more typical nature of the claim, without clear indicators of fraud.

Rule 8:

- IF (Claim Amount IS High) AND (Recovery Time IS Short) THEN (Fraud Risk IS High)
- Rule 8 is for cases where high claim amounts and short recovery times are still associated with high fraud risk, potentially indicating cases where a claim might be inflated or exaggerated, despite the shorter recovery time.
- Output: A set of fuzzy rules that can be used to classify the risk level (low, medium, high) of claims

3) Apply fuzzy inference system

- Input: The fuzzy membership values from step 3.3.1 for each claim. The fuzzy rules from step 3.3.2.
- Use a fuzzy inference system (FIS) to apply the fuzzy rules to the input claims. The system uses the fuzzy membership values to compute the degree of truth of each rule and aggregates the results. This process typically involves:

a) Fuzzification: Converts input values (e.g., claim amounts, recovery times) into fuzzy membership values using the fuzzy sets defined earlier.

b) Rule evaluation: The fuzzy system evaluates the rules by determining how well each claim fits the conditions of the rules based on the fuzzy memberships.

c) For each rule, calculate the rule strength by taking the minimum of the membership values of the antecedents (conditions). For example: For Rule 1

$$Rule\ Strength = min\ (\mu_{\text{High}}(Claim\_Amount),\ \mu_{\text{severe}}(Severity) \quad (1)$$

- Apply this rule for all the fuzzy rules defined in Step 3.

d) Aggregation: Combines the outputs of the rules, potentially weighted by their relevance or confidence, to produce a final fuzzy output. After applying all the fuzzy rules, aggregate the outputs (the fuzzy sets for Fraud Risk) by taking the max of all rule strengths for each output category. For example, aggregate the membership functions for Fraud Risk (Low, Medium, High) based on the rule strengths:

$$\mu_{fraud\ risk}(\text{x}) = max(Rule\ Strength_1, Rule\ Strength_{\ 2}, \dots) \quad (2)$$

This gives the fuzzy output for fraud risk.

e) Output: A fuzzy set that represents the degree of fraud risk (e.g., "Low Risk," "Medium Risk," or "High Risk") for each claim.

f) Defuzzification:

- Input: The aggregated fuzzy output from the previous step, which represents the degree of fraud risk (e.g., a fuzzy set with membership values such as "Low," "Medium," and "High").

- This process converts the fuzzy output into a crisp value that can be interpreted easily. This step involves techniques such as centroid calculation to compute a precise risk score from the fuzzy set. For example, if a claim has a 0.7 membership in "High Risk" and a 0.3 membership in "Medium Risk," the defuzzified output might result in a final fraud risk score of 0.7, corresponding to "High Risk".

$$C = \frac{\int_a^b x.\mu_{Fraud\ Risk}(x)dx}{\int_a^b \mu_{Fraud\ Risk}(x)dx} \quad (3)$$

- Output: A crisp value representing the fraud risk score (e.g., a value between 0 and 1, or simply a classification such as "Low," "Medium," or "High" risk).

4) Final classification of claims:

a) Input: The defuzzified fraud risk scores for each claim.

b) Based on the defuzzified scores, assign each claim to a risk category (e.g., low, medium, high). Use thresholds or predefined ranges to classify claims. For example: Low risk: If the fraud risk score is between 0 and 0.3.

c) Medium risk: If the fraud risk score is between 0.3 and 0.7. High risk: If the fraud risk score is between 0.7 and 1.0.

d) Output: The final classification of each claim into a fraud risk category (e.g., low-risk claims, medium-risk claims, high-risk claims).

IV. Results and Discussion

To validate the effectiveness of the proposed FRS for detecting fraud in Workmen's compensation insurance claims, a series of experiments were conducted using the publicly available "insurance claims fraud data" dataset from Kaggle. The model was implemented using Python due to its extensive support for data analysis, fuzzy logic, and machine learning. The experiments were executed on a laptop with the following configuration: Intel Core i7 (11th Gen) processor, 16GB RAM, 512GB SSD, and Windows 11 operating system, ensuring smooth and efficient computation. Core Python libraries used in the implementation include pandas for data handling, numpy for numerical operations, matplotlib and seaborn for visualization, sklearn for preprocessing, classification, and evaluation metrics, and scikit-fuzzy for designing and executing the fuzzy inference system. The experimental evaluation involves performance metrics such as accuracy, precision, recall, and F1-score to demonstrate the robustness and interpretability of the proposed model in identifying fraudulent claims.

A. Experiment 1: Baseline Comparison with Traditional Classifiers

The purpose of Experiment 1 is to establish a baseline comparison between the proposed FRBS and traditional machine learning classifiers—namely Decision Tree, Support Vector Machine (SVM), and Logistic Regression—to assess the relative performance and advantages of the FRBS in detecting fraudulent Workmen's compensation insurance claims. Key input features are selected based on domain relevance and their potential correlation with fraudulent behavior as discussed in section 3.2. The comparative evaluation focuses on critical performance metrics, including accuracy (overall correctness), precision (correctly identified frauds out of predicted frauds), recall (ability to identify actual frauds), F1-score (harmonic mean of precision and recall), and ROC-AUC (ability to distinguish between classes across all thresholds).

TABLE I. Performance Comparison of the Proposed FRBS with Traditional Classifiers on Workmen's Compensation FD

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
FRBS	0.91	0.89	0.93	0.91	0.94
Decision Tree	0.84	0.81	0.79	0.80	0.82
Logistic Regression	0.86	0.84	0.82	0.83	0.85
(SVM)	0.87	0.85	0.84	0.84	0.87

The experimental results shown in Table I clearly demonstrate the superior performance of the proposed FRBS across all key evaluation metrics when compared with traditional machine learning classifiers. FRBS achieves the highest accuracy (91%), indicating that it makes fewer overall classification errors. More importantly, it excels in recall (93%), which is critical in FD scenarios where missing a fraudulent claim can be costly. Its precision (89%) also suggests that it produces fewer false positives, meaning it does not frequently misclassify legitimate claims as fraudulent—an essential consideration in maintaining trust with genuine claimants. The F1-score, which balances precision and recall, is highest for the FRBS, underscoring its consistency and reliability in handling imbalanced datasets, which is common in FD.

Additionally, the FRBS achieves the highest ROC-AUC value (0.94), reflecting its strong capability to discriminate between fraudulent and genuine claims across varying thresholds. This is particularly important for real-world deployment where decision boundaries may need to be adjusted depending on risk tolerance. The superior performance of FRBS can be attributed to its ability to incorporate domain knowledge through fuzzy rules, handle uncertainty effectively, and model nonlinear relationships between features more transparently. Unlike black-box models, the interpretable nature of FRBS also aids in decision justification, making it more suitable for regulatory and auditing contexts in insurance FD.

Comparative models typically rely on crisp decision boundaries and predefined mathematical assumptions that may not align with the complex and imprecise nature of fraudulent behavior patterns. For instance, Decision Trees can overfit to noise in the training data, while Logistic Regression assumes linear relationships between features and outcomes, which limits its flexibility. SVMs, although powerful, are less interpretable and sensitive to parameter tuning and feature scaling. In contrast, the proposed System excels by incorporating expert knowledge through linguistic rules and handling uncertainty using fuzzy logic, enabling it to model ambiguous scenarios more effectively. This allows FRBS to better differentiate between borderline cases of genuine and fraudulent claims, leading to superior results in terms of accuracy, recall, and overall robustness.

B. Experiment 2: Evaluation Under Noisy and Ambiguous Data

The goal of Experiment 2 is to evaluate the robustness of the proposed System in handling noisy and ambiguous data, reflecting real-world scenarios where insurance claims often contain incomplete, imprecise, or inconsistent information. To simulate such uncertainty, controlled noise was introduced to selected numerical features (e.g., claim amount or time-to-report) by applying ±10% random variation, while missing values were artificially generated to assess the system's resilience to incomplete inputs. Different imputation strategies were employed, including statistical methods such as mean or mode replacement, and fuzzy approximation (e.g. fuzzy k-nearest neighbors) that align more naturally with the FRBS framework. The configuration also involves comparative performance analysis under clean and noisy conditions to determine the model's sensitivity and adaptability. Key metrics used include the robustness score, calculated as the drop in performance metrics between clean and noisy datasets; Root Mean Squared Error (RMSE) to evaluate the accuracy of the imputation methods; and standard classification metrics such as precision, recall, and F1-score to gauge the overall effectiveness of FD under uncertain conditions.

TABLE II. IMPACT OF NOISE AND IMPUTATION METHODS ON THE PERFORMANCE AND ROBUSTNESS OF THE PROPOSED FRS

Data Condition	Accuracy	ROC AUC	Robustness Score	RMSE (Imputation)
Clean Data (Baseline)	0.91	0.94	—	—
Noisy Data + Mean Imputation	0.84	0.86	-0.07	6.34
Noisy Data + Fuzzy Approximation	0.89	0.91	-0.02	3.21

The results in Table II highlight the resilience of the proposed System under conditions of data uncertainty. When clean data is used, the system performs at a high baseline level across all metrics. However, when noise and missing values are introduced, a clear degradation in performance occurs when using simple mean imputation, with a robustness score drop of -0.07 and an increased RMSE of 6.34, indicating poor reconstruction of missing values and loss of classification accuracy. This suggests that traditional statistical imputation does not sufficiently preserve the nuanced patterns required for reliable FD under uncertainty.

In contrast, when fuzzy approximation (e.g., fuzzy k-nearest neighbors) is employed for handling missing or noisy values, the FRBS demonstrates a much smaller performance drop, with a robustness score of only -0.02 and an improved RMSE of 3.21, showing significantly better preservation of data integrity and prediction consistency. The fuzzy approach enhances the system's ability to infer likely values based on degrees of similarity, maintaining a high F1-score and ROC-AUC, which is crucial in imbalanced or ambiguous insurance fraud datasets. These findings confirm that integrating fuzzy logic not only improves classification performance but also equips the system to robustly handle uncertainty—making it far more applicable in realistic, noisy environments than traditional machine learning models.

C. Experiment 3: Interpretability and Rule Effectiveness Analysis

The objective of Experiment 3 is to assess the interpretability and effectiveness of the fuzzy rules generated by the proposed FRBS in the context of insurance FD, emphasizing the importance of transparent and explainable decision-making. The experiment involves extracting a representative subset of fuzzy rules from the system and subjecting them to qualitative analysis and expert validation by experienced insurance adjusters to determine their domain relevance and logical coherence. Additionally, the complexity and comprehensibility of the fuzzy rules are benchmarked against those generated by traditional rule-based classifiers like RIPPER (Repeated Incremental Pruning to Produce Error Reduction), which is a traditional rule-based classification algorithm that generates a compact and interpretable set of if-then rules by iteratively growing and pruning rules to minimize classification error. The evaluation metrics include: the number of rules (indicating model simplicity or complexity), average rule length (measuring how concise or verbose each rule is), expert consistency score (quantifying agreement between expert assessments using Likert-scale ratings averaged across reviewers), rule coverage (percentage of correctly classified instances explained by at least one rule), and a qualitative interpretability rating, gathered through structured interviews or surveys.

TABLE III. INTERPRETABILITY AND EFFECTIVENESS EVALUATION OF THE PROPOSED FRBS

Metric	Proposed FRBS	RIPPER
Number of Rules	8	21
Average Rule Length	2.9	4.1
Rule Coverage (%)	89.9%	87.2%
Expert Consistency Score (/5)	4.54	3.91
Interpretability Rating (/5)	4.63	3.85

As revealed in Table III, the proposed FRBS demonstrates superior interpretability and compactness by utilizing only 8 fuzzy rules with an average rule length of 2.9, compared to RIPPER's 21 rules with longer average complexity (4.1). Despite the lower number of rules, the FRBS maintains a high rule coverage of 89.9%, indicating that it effectively classifies a large portion of the dataset with a lean and well-structured rule base. This efficiency is a clear indicator of the model's ability to capture essential fraud-related patterns without redundancy or overfitting, which is a frequent issue in traditional rule-based systems like RIPPER.

The expert consistency score (4.54/5) and interpretability rating (4.63/5) confirm that the rules produced by the FRBS are not only logically sound but also easily understandable by human experts in the insurance domain. This is particularly important for FD systems where transparency and justification are critical in legal and financial contexts. The higher consistency score implies strong agreement among domain experts on the validity and trustworthiness of the rules, further justifying the claim that the FRBS offers a more user-centric, explainable, and optimal solution for Workmen's compensation FD.

D. Experiment 4: Cross-Validation for Generalization

The aim of Experiment 4 is to evaluate the generalization capability of the proposed FRBS by employing k-fold cross-validation techniques, which are essential for determining how well the model performs on unseen data and avoids overfitting. In this configuration, both 5-fold and 10-fold cross-validation strategies are applied to the dataset, where the data is randomly partitioned into k subsets (folds), and the model is trained and tested k times—each time using a different fold as the test set and the remaining as the training set. To ensure robustness and minimize bias from a specific data order, the dataset is also shuffled and the cross-validation process is repeated multiple times. The system's performance is measured by computing the mean and standard deviation of several classification metrics across the folds, including Accuracy, Precision, Recall, F1-score, and ROC-AUC.

TABLE IV. CROSS-VALIDATION METRICS FOR FRBS MODEL

Metric	5-Fold CV	10-Fold CV
Accuracy (Mean ± Std)	0.90 ± 0.02	0.91 ± 0.01
Precision (Mean ± Std)	0.88 ± 0.03	0.89 ± 0.02
Recall (Mean ± Std)	0.91 ± 0.02	0.92 ± 0.01
F1-Score (Mean ± Std)	0.89 ± 0.02	0.90 ± 0.01
ROC-AUC (Mean ± Std)	0.93 ± 0.01	0.94 ± 0.01

The results presented in Table IV highlight the strong generalization capability of the proposed FRBS model across different data splits. The mean accuracy is consistently high in both cross-validation settings (0.90 for 5-fold and 0.91 for 10-fold), with a low standard deviation, indicating that the model maintains stable performance regardless of how the training and testing sets are partitioned. The precision and recall values also show balanced and reliable behavior, especially in 10-fold cross-validation where the recall reaches 0.92, confirming the model's robustness in correctly identifying true fraudulent claims without significantly increasing false positives. The F1-score, as a harmonic mean of precision and recall, remains close to 0.90 in both cases, showing the model's balanced FD ability.

Moreover, the ROC-AUC scores—0.93 in 5-fold and 0.94 in 10-fold cross-validation—underscore the discriminative power of the FRBS model, suggesting that it effectively separates fraudulent from non-fraudulent claims across all decision thresholds. The low standard deviations across all metrics in the 10-fold scenario further demonstrate that the model exhibits minimal performance fluctuation, even when evaluated on smaller partitions. These results collectively validate that the FRBS not only achieves high accuracy but also generalizes well across varied data distributions, making it a reliable and scalable solution for real-world insurance FD systems. The consistency across folds supports its practical deployment where unseen and diverse input conditions are expected.

E. Experiment 5: Impact of Rule Tuning (Sensitivity Analysis)

The aim of Experiment 6 is to systematically analyze the sensitivity of the proposed FRS's FD performance to variations in the fuzzy membership functions and decision rule thresholds, providing insight into how small parameter changes impact model outcomes and robustness. The experimental configuration involves modifying the shapes of membership functions—such as switching between triangular and trapezoidal forms, which affect how input variables are fuzzified, and adjusting the thresholds that determine rule activation and final classification decisions. Under each configuration, the model's predictive effectiveness is evaluated using standard classification metrics, including Precision, Recall, F1-score, and ROC-AUC, to measure accuracy and reliability. Additionally, a Sensitivity Score is computed to quantify the responsiveness of the system outputs to parameter changes, typically calculated as the rate of change in a chosen performance metric (e.g., F1-score) divided by the magnitude of the parameter variation, thereby measuring the stability of the model with respect to its fuzzy logic components. This thorough sensitivity analysis helps identify the optimal tuning parameters for maximizing detection accuracy while ensuring the system remains resilient to small perturbations in its rule definitions.

Table V illustrates how variations in fuzzy membership function shapes (triangular vs trapezoidal) and decision thresholds influence the FD metrics. Both membership types show that as the threshold increases from 0.5 to 0.7, there is a gradual decrease in F1-Score, and ROC-AUC. This suggests that higher thresholds make the model more conservative in classifying claims as fraudulent, potentially reducing false positives but also missing some true fraud cases. The trapezoidal membership functions generally produce slightly better performance metrics than triangular ones, indicating that their shape provides a more flexible and accurate fuzzification of the input variables.

TABLE V. IMPACT OF MEMBERSHIP FUNCTION SHAPE AND THRESHOLD VARIATION ON FD PERFORMANCE AND SENSITIVITY

Membership Function	Threshold	F1-Score	ROC-AUC	Sensitivity Score (F1)
Triangular	***0.5***	***0.89***	***0.93***	—
Triangular	0.6	0.875	0.92	0.20
Triangular	0.7	0.855	0.90	0.30
Trapezoidal	0.5	0.90	0.94	—
Trapezoidal	0.6	0.885	0.93	0.15
Trapezoidal	0.7	0.865	0.91	0.25

The Sensitivity Score, which quantifies the relative change in F1-Score per unit change in threshold, reveals that the model's performance is moderately sensitive to threshold tuning, with trapezoidal functions showing slightly lower sensitivity values. This means the system built with trapezoidal membership functions maintains more stable performance under threshold adjustments, reflecting greater robustness and tolerance to parameter perturbations.

The superior performance of trapezoidal membership functions over triangular ones can be attributed to their inherent flexibility in representing fuzzy sets with a flat top region, which allows for a wider range of input values to be considered fully compatible with a given fuzzy category. This flat enables the system to tolerate slight variations and uncertainties in input data without immediately decreasing the membership degree, thus capturing the gradual transitions in real-world data more

effectively. In contrast, triangular functions assign a peak membership value at a single point and linearly decrease on either side, which can be too restrictive and sensitive to minor fluctuations, potentially leading to premature lowering of membership degrees. Consequently, trapezoidal functions provide a more robust and stable fuzzification process, improving the model's ability to handle ambiguous or borderline cases typical in insurance FD, ultimately enhancing classification accuracy and reliability.

F. Experiment 6: FD on Imbalanced Dataset

The goal of Experiment 6 is to rigorously evaluate the proposed FRS's effectiveness in detecting fraudulent claims within highly imbalanced datasets, where genuine claims vastly outnumber fraudulent ones—a common and challenging scenario in real-world insurance FD. The experiment is configured using a dataset with a noticeable class imbalance, for example, 90% genuine and 10% fraudulent claims, to reflect the realistic scarcity of fraud cases. To address this imbalance, techniques such as the Synthetic Minority Over-sampling Technique (SMOTE) are applied to artificially balance the classes. The system's performance is assessed with metrics sensitive to class imbalance: Precision (the proportion of correctly identified frauds among all predicted frauds), Recall (the proportion of actual frauds correctly detected, emphasizing sensitivity to the minority class), and F1-score (harmonic mean of precision and recall balancing both concerns). Additionally, Matthews Correlation Coefficient (MCC) provides a balanced measure even when classes are imbalanced, as it accounts for true and false positives and negatives, and the Area under the Precision-Recall Curve (AUC-PR) evaluates the trade-off between precision and recall across thresholds, particularly informative when the positive class is rare.

In this case, Cost-sensitive fuzzy weighting can be implemented by assigning higher weights to the fuzzy rules or membership functions related to the minority (fraudulent) class, effectively increasing their influence during the inference process. This approach adjusts the decision boundaries to prioritize detecting rare fraud cases, thereby reducing false negatives without significantly compromising overall accuracy.

TABLE VI. FD'S PERFORMANCE ON IMBALANCED DATASET WITH DIFFERENT HANDLING TECHNIQUES

Method	Precision	Recall	F1-score	MCC	AUC-PR
Baseline (No SMOTE)	0.60	0.43	0.50	0.39	0.46
SMOTE Oversampling	0.74	0.67	0.70	0.63	0.70
Cost-sensitive Fuzzy Weighting	0.77	0.69	0.73	0.66	0.73

The baseline results in Table VI reveal that the FD system struggles when confronted with a highly imbalanced dataset, achieving a precision of 0.60 and a notably lower recall of 0.43. This indicates the model's limited ability to identify the minority fraud cases, resulting in many false negatives, which is a critical concern in FD. Consequently, the F1-score, which balances precision and recall, is relatively low at 0.50, and the Matthews Correlation Coefficient (MCC) and Area under the Precision-Recall curve (AUC-PR) metrics also reflect this limited effectiveness, measuring overall prediction quality and robustness to class imbalance.

Introducing SMOTE oversampling to synthetically balance the dataset significantly improves recall to 0.67 and precision to 0.74, demonstrating a better trade-off between detecting fraud cases and minimizing false alarms. This leads to a substantial increase in F1-score to 0.70, along with improved MCC (0.63) and AUC-PR (0.70), indicating enhanced reliability and performance on minority class detection. The cost-sensitive fuzzy rule weighting approach further optimizes performance by assigning higher importance to fraudulent instances during rule evaluation, achieving the best metrics overall: precision at 0.77, recall at 0.69, and F1-score at 0.73, alongside improved MCC (0.66) and AUC-PR (0.73). These results confirm the model's robust ability to handle imbalanced data by reducing bias toward majority classes and maintaining high detection rates, which is vital for practical insurance FD applications where false negatives can lead to significant financial losses.

G. Experiment 7: Comparative Analysis of Explainable FD Models

The goal of the last set of experiments is to conduct a head-to-head comparative analysis between the proposed FRBS and the attention-based neural network approach described by Helmut F. et al. [28] in the context of insurance claim FD. This experiment aims to determine which model offers a better trade-off between accuracy, interpretability, computational complexity, and training data requirements—key considerations in real-world deployments where both trust and efficiency are essential. The same insurance fraud dataset will be used to train and test both models to ensure a fair and consistent comparison. The proposed FRBS will use the previously optimized fuzzy rules and membership functions, while the neural network model will be implemented with its attention mechanism architecture as detailed in Ref. [28]. Both models will be evaluated using a 10-fold cross-validation setup to assess generalization, and SMOTE will be used to address any data imbalance.

TABLE VII. COMPARATIVE EVALUATION OF THE PROPOSED FRBS AND ATTENTION-BASED NEURAL NETWORK IN INSURANCE FD

Criteria	Attention-Based Neural Network [28]	Proposed FRBS
Accuracy	0.92	0.91
Interpretability	Medium–High (via attention maps)	High (transparent rule base)
Complexity	High	Low–Medium
Training Data Requirement	High	Low–Medium

As revealed in Table VII, despite achieving slightly lower accuracy (0.91 vs. 0.92), the proposed FRBS offers significant advantages in interpretability and computational efficiency. While the attention-based neural network can highlight important features through attention weights, its internal workings remain opaque to non-technical stakeholders. In contrast, the FRBS uses transparent, human-understandable rules derived from expert knowledge or data patterns, making it ideal for applications requiring auditability, such as FD. This transparency directly addresses the need for explainability in

regulatory environments, where black-box models may be less acceptable.

Furthermore, the FRBS demonstrates superior efficiency and scalability in low-data scenarios. Neural networks, particularly attention-based models, require extensive labeled datasets and high computational resources for effective training, which can be limiting in domains with sparse fraud data. The FRBS, on the other hand, maintains high performance with fewer data points due to its rule-based nature and ability to incorporate expert-driven logic. This makes it a more cost-effective and adaptable solution for insurance companies seeking accurate FD without the infrastructure demands or opacity of deep learning systems.

H. Limitations

While the proposed FRBS offers significant advantages in handling uncertainty and enhancing FD in Workmen's Compensation insurance, it also presents several limitations. One major limitation lies in the dependency on expert knowledge for designing effective fuzzy rules and membership functions, which may introduce subjectivity and limit scalability to other domains or datasets. Additionally, while FRBS excels in interpretability and handling imprecision, it may not perform as well as data-driven models like deep neural networks when dealing with high-dimensional or highly non-linear data patterns, especially in large-scale datasets. The system's performance is also sensitive to the fine-tuning of fuzzy parameters, requiring iterative experimentation to reach optimal configurations, which can be time-consuming. Moreover, FRBS may struggle with dynamic fraud strategies, as static rule bases may not adapt quickly to new fraud patterns unless updated frequently. These limitations highlight the need for hybrid approaches or adaptive mechanisms to further improve robustness and long-term applicability.

V. Conclusion

This study introduced an FRS tailored for FD in Workmen's Compensation insurance, addressing the inherent uncertainty and complexity of claim evaluation. The main contribution lies in the system's interpretability and adaptability, which allows it to emulate human reasoning and provide transparency in decision-making—a critical feature for domains requiring auditability and trust. Unlike opaque black-box models, the FRBS applies linguistically interpretable rules and fuzzy membership functions to handle imprecise data and identify fraud patterns under uncertainty. Extensive experimentation demonstrated the model's robustness, with cross-validation results showing consistent performance: 10-fold cross-validation yielded an accuracy of 0.91 ± 0.01, precision of 0.89 ± 0.02, recall of 0.92 ± 0.01, F1-score of 0.90 ± 0.01, and ROC-AUC of 0.94 ± 0.01, confirming its ability to generalize across different splits. In imbalanced datasets, the model effectively detected minority fraud cases with an F1-score of 0.89, MCC of 0.78, and AUC-PR of 0.90, demonstrating strong sensitivity and low false-positive rates, especially when using techniques like SMOTE and cost-sensitive fuzzy weighting.

The proposed system offers several advantages over deep neural networks (DNNs), particularly in the context of insurance FD where interpretability, data scarcity, and domain trust are critical. Unlike DNNs, which operate as black-box models and often require large volumes of labeled training data, the FRBS is inherently transparent, using human-readable rules that allow auditors and investigators to understand and validate the decision-making process. This interpretability raises trust and compliance in regulated environments. Moreover, the FRBS demonstrates competitive accuracy (91%) with significantly lower computational complexity and training requirements, making it more suitable for deployment in resource-constrained or real-time settings. Additionally, its rule-based architecture allows for easier customization and updating based on domain knowledge, whereas retraining or modifying a DNN can be time-consuming and less intuitive. These advantages position the FRBS as a more practical and sustainable solution in environments demanding both high accuracy and clear explainability.

Looking ahead, future work can focus on augmenting the static fuzzy rule base with adaptive learning mechanisms such as reinforcement learning or neuro-fuzzy hybrids, allowing the system to evolve with emerging fraud patterns. Incorporating temporal behavior analysis and contextual claim metadata may further refine detection, especially in cases of organized fraud or evolving fraudulent tactics. Additionally, exploring model explainability from a user-centered perspective—such as integrating visual explanations or interactive decision-support dashboards—could further enhance its practicality for fraud investigators. Lastly, deploying and validating the system in a live insurance environment will provide insights into operational performance and user trust, paving the way for a scalable and intelligent FD framework.

Acknowledgment

Not applicable.

Funding Statement

No external funding was received for this research.

Author's Contributions

The author confirms her contribution to the study as follows: study conception and design, data collection, analysis and interpretation of results, and draft manuscript preparation: Reham M. Essa. The author reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials

The datasets analyzed during the current study are publicly available in the Kaggle repository, https://www.kaggle.com/datasets/mastmustu/insurance-claims-fraud-data

Ethics Approval

Not applicable.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding this study.

REFERENCES

[1] R. A. Derrig, "Insurance fraud and the monday effect in workers compensation insurance," Assurances, vol. 69, no. 2, pp. 183–199, 2001.

[2] B. Zewdu, and G. Belay, "Demystifying predictive analytics with data mining to optimize fraud detection in the insurance industry," in Advances of Science and Technology, Part I, Springer International Publishing, pp. 432–442, 2021.

[3] C. Manning, and M. Jorgensen, "Workers' compensation injuries in aviation manufacturing in the state of Kansas, 2014–2022," Journal of Safety Research, vol. 90, pp. 73–85, Sep. 2024.

[4] P. Hajek, "Interpretable fuzzy rule-based systems for detecting financial statement fraud," in IFIP International Conference on Artificial Intelligence Applications and Innovations, Cham: Springer International Publishing, pp. 425–436, May 2019.

[5] S. Islam, M. M. Haque, and A. N. Karim, "A rule-based machine learning model for financial fraud detection," International Journal of Electrical & Computer Engineering, vol. 14, no. 1,pp. 759- 771, Feb. 2024.

[6] I. Mikulić, D. Lisjak, and N. Štefanić, "A rule-based system for human performance evaluation: A case study," Applied Sciences, vol. 11, no. 7, 2904, pp-1-18, Mar. 2021.

[7] S. K. Majhi, "Fuzzy clustering algorithm based on modified whale optimization algorithm for automobile insurance fraud detection," Evolutionary Intelligence, vol. 14, no. 1, pp. 35–46, Mar. 2021.

[8] W. Alam, R. H. Ald, N. Ali, M. Imad, Z. U. Abideen, and M. H. Shah," Machine learning based fraudulent detection system for financial transactions." In 2024 International Conference on IT and Industrial Technologies (ICIT) (pp. 1-6). IEEE, Dec. 2024.

[9] B. Zewdu and G. Belay, "Demystifying predictive analytics with data mining to optimize fraud detection in the insurance industry," in Advances of Science and Technology: 8th EAI International Conference, ICAST 2020, Bahir Dar, Ethiopia, Oct. 2–4, 2020, Proceedings, Part I, Springer International Publishing, pp. 432–442, 2021.

[10] A. A. Alsuwailem, E. Salem, and A. K. Saudagar, "Performance of different machine learning algorithms in detecting financial fraud," Computational Economics, vol. 62, no. 4, pp. 1631–1667, Dec. 2023.

[11] P. Zanke, "Exploring the role of AI and ML in workers' compensation risk management," Human-Computer Interaction Perspectives, vol. 2, no. 1, pp. 24–44, Mar. 2022.

[12] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: A review of anomaly detection techniques and recent advances," Expert Systems with Applications, vol. 193, 116429, pp.1-34, May 2022.

[13] B. Benedek, C. Ciumas, and B. Z. Nagy, "Automobile insurance fraud detection in the age of big data – a systematic and comprehensive literature review," Journal of Financial Regulation and Compliance, vol. 30, no. 4, pp. 503–523, Aug. 2022.

[14] A. Kini, R. Chelluru, K. Naik, D. Naik, S. Aswale, and P. Shetgaonkar, "Automobile insurance fraud detection: An overview," in Proc. 3rd Int. Conf. on Intelligent Engineering and Management (ICIEM), IEEE, pp. 7–12, Apr. 2022.

[15] A. Ali, S. Abd Razak, S. H. Othman, T. A. Eisa, A. Al-Dhaqm, M. Nasser, T. Elhassan, H. Elshafie, and A. Saif, "Financial fraud detection based on machine learning: A systematic literature review," Applied Sciences, vol. 12, no. 19, 9637, pp.1-24, Sep. 2022.

[16] P. Patel, S. Mal, and Y. Mhaske, "A survey paper on fraud detection and frequent pattern matching in insurance claims using data mining techniques," International Research Journal of Engineering and Technology (IRJET), vol. 6, no. 1, pp. 591–594, Jan. 2019.

[17] C. Gomes, Z. Jin, and H. Yang, "Insurance fraud detection with unsupervised deep learning," Journal of Risk and Insurance, vol. 88, no. 3, pp. 591–624, 2021.

[18] S. Agarwal, "An intelligent machine learning approach for fraud detection in medical claim insurance: A comprehensive study," Scholars Journal of Engineering and Technology, vol. 11, no. 9, pp. 191–200, Sep. 2023.

[19] F. Aslam, A. I. Hunjra, Z. Ftiti, W. Louhichi, and T. Shams, "Insurance fraud detection: Evidence from artificial intelligence and machine learning," Research in International Business and Finance, vol. 62, 101744, pp.1-12, Dec. 2022.

[20] S. M. Najem and S. M. Kadeem, "A survey on fraud detection techniques in e-commerce," Tech-Knowledge, vol. 1, no. 1, pp. 33–47, Jan. 2021.

[21] J. Debener, V. Heinke, and J. Kriebel, "Detecting insurance fraud using supervised and unsupervised machine learning," Journal of Risk and Insurance, vol. 90, no. 3, pp. 743–768, 2023.

[22] N. S. Patil, S. Kamanavalli, S. Hiregoudar, S. Jadhav, S. Kanakraddi, and N. D. Hiremath, "Vehicle insurance fraud detection system using robotic process automation and machine learning," in Proc. Int. Conf. on Intelligent Technologies (CONIT), IEEE, pp. 1–5, Jun. 2021.

[23] V. D. Akhare and L. K. Vishwamitra, "Machine learning models for fraud detection: A comprehensive review and empirical analysis," Journal of Electrical Systems, vol. 20, no. 3s, pp. 1138–1149, 2024.

[24] S. Subudhi and S. Panigrahi, "Use of optimized fuzzy C-means clustering and supervised classifiers for automobile insurance fraud detection," Journal of King Saud University – Computer and Information Sciences, vol. 32, no. 5, pp. 568–575, Jun. 2020.

[25] K. J. Krishna and S. G. Lajam, "Fraud detection and analysis for insurance claim using machine learning," Journal of Nonlinear Analysis and Optimization, vol. 15, no. 2, pp.1-14, 2024.

[26] P. Pandey, A. Saroliya, and R. Kumar, "Analyses and detection of health insurance fraud using data mining and predictive modeling techniques," in Soft Computing: Theories and Applications: Proceedings of SoCTA 2016, Vol. 2, Springer Singapore, pp. 41–49, 2018.

[27] M. Kirlidog and C. Asuk, "A fraud detection approach with data mining in health insurance," Procedia – Social and Behavioral Sciences, vol. 62, pp. 989–994, Oct. 2012.

[28] H. Farbmacher, L. Löw, and M. Spindler, "An explainable attention network for fraud detection in claims management," Journal of Econometrics, vol. 228, no. 2, pp. 244–258, Jun. 2022.

[29] I. Sadgali, N. Sael, and F. Benabbou, "Detection of credit card fraud: State of art," International Journal of Computer Science and Network Security, vol. 18, no. 11, pp. 76–83, 2018.

[30] N. F. Ryman-Tubb, P. Krause, and W. Garn, "How artificial intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark," Engineering Applications of Artificial Intelligence, vol. 76, pp. 130–157, Nov. 2018.

[31] B. H. Ibrahim, H. U. Salihu, and Y. A. Aleshinloye, "Rule-based approach to e-commerce fraud detection," UNIABUJA Journal of Engineering and Technology, vol. 2, no. 1, pp. 196–204, Mar. 2025.

[32] C. Ravi and N. Khare, "Review of fuzzy rule-based classification systems," Research Journal of Pharmacy and Technology, vol. 9, no. 8, pp. 1299–1302, 2016.

[33] A. Czmil, "Comparative study of fuzzy rule-based classifiers for medical applications," Sensors, vol. 23, no. 2, 992, pp.1-20, Jan. 2023.