# QUANTUM WOLF

DATA INTELLIGENCE RESEARCH LAB

DATE:14-02-2025 to 19-02-2025

NAME: ABDUL ZUHAIL M

**Fraud Detection System for Loan Transactions**

**1. Problem Statement**

The financial sector faces increasing challenges in detecting fraudulent loan transactions. Fraudulent activities such as misrepresentation of financial data, fake loan applications, and deliberate repayment defaults lead to major financial losses. Traditional fraud detection methods are:

* **Manual –** Slow and requires human intervention.
* **Error-Prone –** Higher chances of oversight and misclassification.
* **Time-Consuming –** Requires extensive effort for verification.

**Need for an Automated System:**

* Accurately detect fraudulent transactions.
* Provide real-time alerts to prevent financial losses.
* Analyze fraud trends and patterns for risk management.

**2. Solution Overview**

The proposed Automated Fraud Detection System using KNIME leverages Machine Learning (Random Forest Algorithm) and rule-based filtering to:

* Preprocess and clean transaction data.
* Train a fraud detection model.
* Predict fraud risk scores for new transactions.
* Generate real-time fraud alerts.
* Provide an interactive dashboard for fraud monitoring.

**3. Working Process**

**Step 1: Data Collection**

* **Source:** Loan transaction datasets.**Method:** CSV Reader Node in KNIME loads the dataset.
* **Output:** A structured dataset containing relevant loan transaction details.

**Step 2: Data Preprocessing**

* **Objective:** Prepare dataset by handling missing values and transforming categorical data.
* **Process:**
  + **Missing Value Handler Node →** Fills missing values (numerical: mean/median, categorical: most frequent).
  + **One-Hot Encoder Node →** Converts categorical values into numerical format.
  + **Normalizer Node →** Scales numerical features for better accuracy.

**Step 3: Model Training & Evaluation**

* **Objective:** Train a fraud detection model and assess performance.
* **Process:**
  + **Partitioning Node →** Splits dataset into training (70%) and testing (30%).
  + **Random Forest Learner Node →** Trains the model using loan attributes (loan amount, credit score, repayment status).
  + **Scorer Node →** Evaluates model using accuracy, precision, recall, and F1-score.
* **Output:** Trained fraud detection model.

**Step 4: Fraud Prediction on New Transactions**

* **Objective**: Predict fraud risk for incoming loan applications.
* **Process:**
  + **Random Forest Predictor Node →** Classifies transactions as fraudulent or genuine.
  + **Math Formula Node →** Computes risk score:
    - (1 - Prediction Confidence) \* Loan Amount
    - **Rule-Based Row Filter Node →** Flags high-risk transactions (Risk Score > 0.2 → Fraudulent).
* **Output**: Fraud risk classification, high-risk transactions flagged for review.

**Step 5: Fraud Alerts & Storage**

* **Objective:** Notify stakeholders about fraud and store flagged transactions.
* **Process:**
  + **Rule-Based Row Filter Node** → Extracts fraudulent transactions.
  + **Send Email Node** → Sends alerts to stakeholders.
  + **CSV Writer Node →** Saves fraud detection results.
* **Output:** Real-time alerts and stored fraud cases for future analysis.

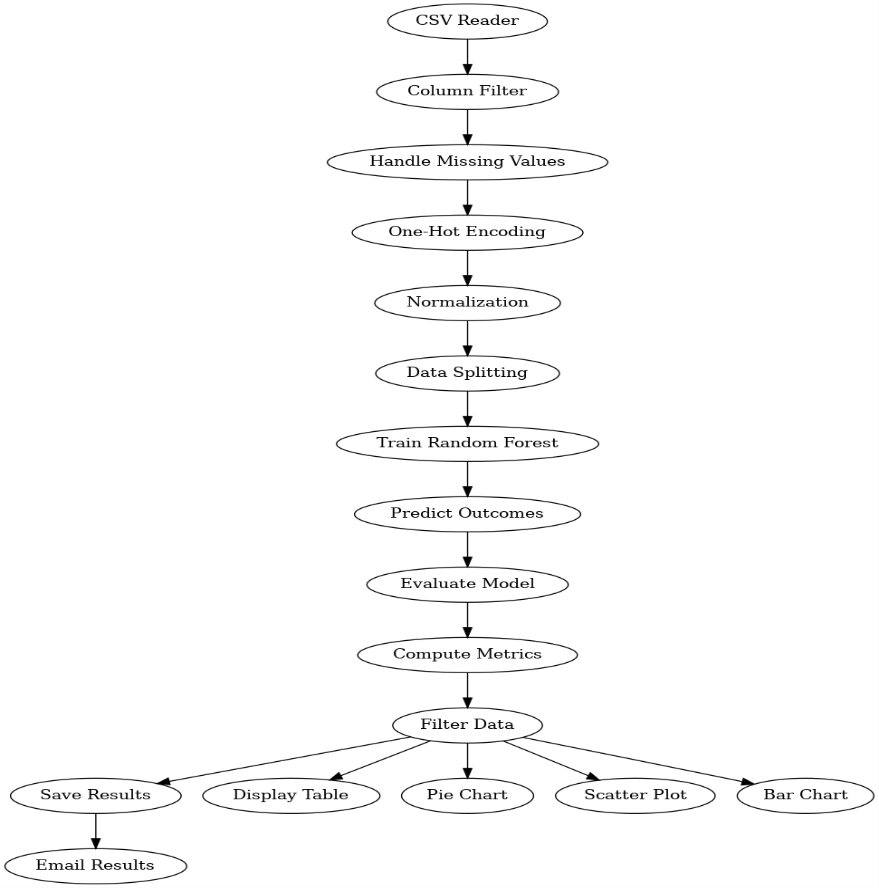
**Step 6: Data Visualization & Dashboard**

* **Objective**: Provide an interactive dashboard for fraud monitoring.
* **Process:**
  + **Table View Node →** Displays structured fraud data.
  + **Visualization Nodes:**
    - **Bar Chart →** Shows fraud trends by loan type and industry.
    - **Pie Chart** → Illustrates fraudulent vs. genuine transactions.
    - **Scatter Plot** → Plots fraud risk scores against loan amounts.
* **Output:** A user-friendly dashboard for fraud monitoring.

**4. Final Output**

* Trained fraud detection model with high accuracy.
* Automated classification of transactions.
* Real-time fraud alerts.
* Dashboard for fraud trend analysis.

**5.Flow Diagram**



**6. Conclusion**

The Fraud Detection System automates fraud identification using Machine Learning & KNIME workflows. It:

* Detects fraudulent transactions efficiently.
* Generates real-time alerts for suspicious activities.
* Provides an interactive monitoring dashboard.

**Future Enhancements:**

* Real-time streaming for instant fraud detection.
* Integration with deep learning models for improved accuracy.