System Model for Fault Detection in Wireless Sensor Networks (WSNs)

1. Objective

Detect various faults in a Wireless Sensor Network (WSN) using machine learning classifiers, specifically focusing on the Random Forest (RF) classifier, which showed superior performance in terms of accuracy and fault detection.

2. Fault Classifiers Used

- Support Vector Machine (SVM): Supervised learning algorithm used to classify faults byhyperplane separation.

- Convolutional Neural Network (CNN): Inspired by biological neural networks to detect and classifyfaults.

- Multilayer Perceptron (MLP): A feed-forward artificial neural network used for fault detection.

- Stochastic Gradient Descent (SGD): Linear model that estimates faults over time.

- Random Forest (RF): Ensemble learning method that uses decision trees to classify faults.

- Probabilistic Neural Network (PNN): Calculates the probability distribution of fault detection.

3. Fault Types

- Offset Fault: Incorrect calibration resulting in an offset in sensor readings.

- Gain Fault: Processing circuit deviates from the actual gain required.

- Stuck-at Fault: A digital circuit is stuck permanently at a zero or one value.

- Out-of-Bounds Fault: Sensor readings exceed the normal range.

- Spike Fault: Sudden spikes in sensor data readings.

- Data Loss Fault: Loss of data due to sensor failure or communication fault.**4. System Model**

1. Data Generation: Simulated sensor data for temperature and humidity with six fault types.

2. Data Preprocessing: Filled missing values from data-loss faults with mean temperature.

3. Model Training: A Random Forest classifier trained on the dataset using temperature and humidity as features.

4. Evaluation: The model achieved an overall accuracy of 84%, with Random Forest showing good performance.
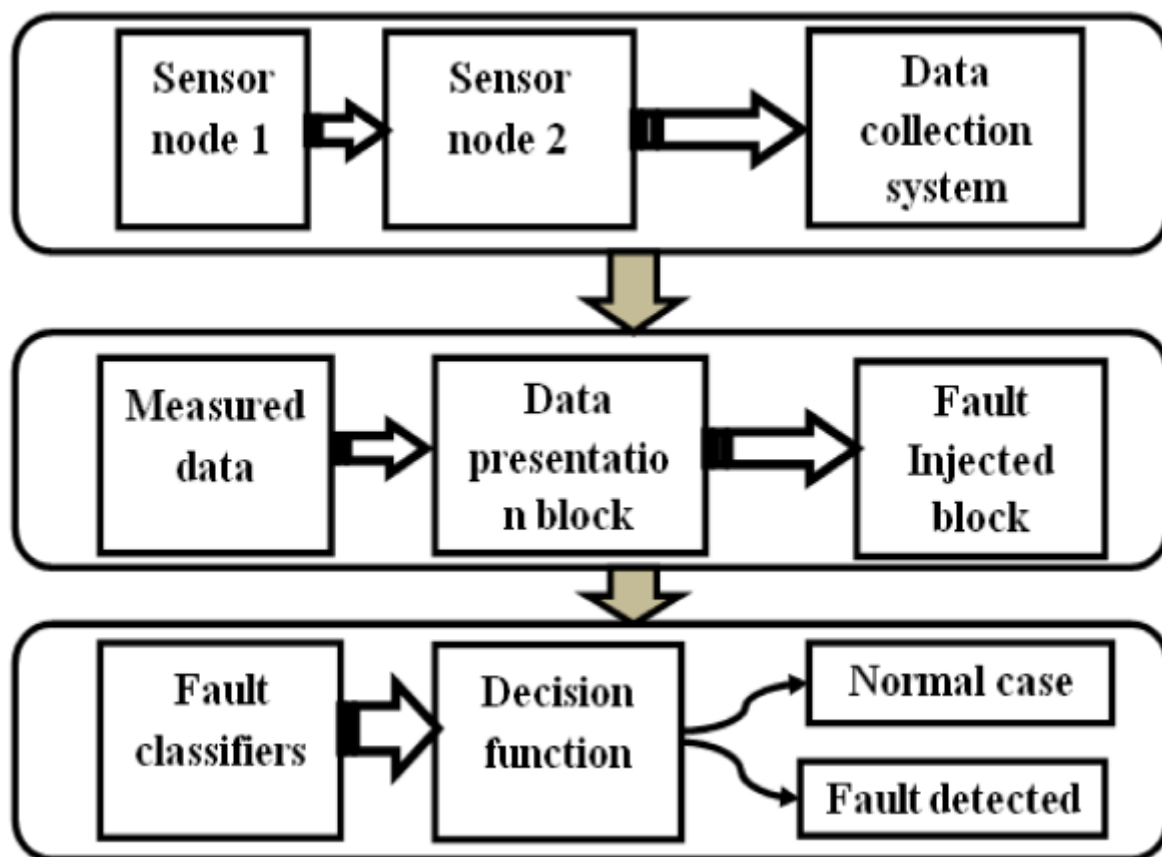
5. Evaluation Metrics

Confusion Matrix:

The confusion matrix visualizes the model's performance across different fault types.

Classification Report:

- Normal: 100% Accuracy

- Gain: 94% Accuracy

- Offset: 63% Accuracy

- Stuck-at: 26% Accuracy

- Out-of-Bounds: 73% Accuracy

- Spike: 100% Accuracy

- Data Loss: 100% Accuracy