Project Report: IoT Fire Alarm Detection System

Student: Halil Berkay Akkaya **Course:** Internet of Things and Applied Data Science
Date: December 2025

1. Executive Summary The objective of this project is to develop a reliable fire detection system using Internet of Things (IoT) sensor data. Traditional smoke detectors often rely on simple threshold mechanisms, leading to false alarms or delayed responses. This project leverages Machine Learning (Random Forest) to analyze multivariate sensor data (Temperature, Humidity, TVOC, eCO2) for accurate fire classification. The final solution is deployed as an interactive dashboard on AWS EC2.

2. Problem Statement

- **Goal:** Classify fire alarm status for indoor environments.
- **Impact:** Enhances resident safety and reduces response time to potential fire hazards.
- **Data Source:** IoT sensor data collected at real-time intervals.
- **Success Metric:** High classification accuracy (>95%) and low false-positive rate.
- **Constraints:** Computational efficiency for potential edge deployment.

3. Data Collection and Processing We utilized the "Smoke Detection Dataset" (Open Source / Kaggle), which consists of over 62,000 readings from a diverse set of IoT sensors.

- **Key Features:** Temperature[C], Humidity[%], TVOC[ppb], eCO2[ppm], Raw H2, Raw Ethanol, Pressure[hPa], PM1.0, PM2.5.
- **Data Cleaning:** Irrelevant columns such as timestamps (UTC) and sample counters (CNT) were removed to prevent data leakage and ensure model generalization.
- **Preprocessing:** The data was checked for null values and split into training (80%) and testing (20%) sets.

4. Exploratory Data Analysis (EDA) Visualization revealed significant correlations between specific gas concentrations and fire events:

- **Heatmap Analysis:** Humidity and TVOC (Volatile Organic Compounds) showed a higher correlation with the "Fire Alarm" status compared to temperature alone.
- **Class Balance:** The dataset contained a higher number of "Fire" instances compared to "No Fire," which was handled during model evaluation.

5. Modeling Methodology A **Random Forest Classifier** was selected for this task due to its robustness against overfitting and ability to handle non-linear relationships in sensor data.

- **Model:** Random Forest (n_estimators=100).
- **Training:** The model was trained on the processed feature set.
- **Performance:** The model achieved an accuracy of ~99% on the test set, demonstrating high reliability in distinguishing between safe and hazardous environments.

6. Cloud Deployment Architecture To demonstrate an end-to-end IoT pipeline, the solution was deployed to the cloud:

- **Platform:** Amazon Web Services (AWS) EC2 (Instance Type: t3.micro/Ubuntu).
- **Interface:** A user-friendly dashboard was built using **Streamlit**, allowing users to simulate sensor inputs and receive real-time fire status predictions.
- **Accessibility:** The application is accessible via a public IP address on port 8501.

7. Conclusion This project successfully demonstrates the application of Data Science in the IoT domain. By integrating sensor data with machine learning, we created a system capable of intelligent fire detection, moving beyond simple hardware thresholds. The successful deployment on AWS EC2 validates the practical applicability of the solution.