AIoT LAB

Experiment 2: Occupancy-Based Smart Home Automation using Unsupervised Learning (K-Means, Hierarchical Clustering)

Aim

To design and implement an AIoT-based smart home automation system using unsupervised machine learning algorithms (K-Means, Hierarchical clustering) on publicly available IoT sensor data and evaluate their performance using accuracy, confusion matrices, and dendrogram visualization.

Apparatus / Software Required

- Python 3.x
- Pandas
- NumPy
- Scikit-learn
- Matplotlib
- SciPy
- UCI Occupancy Detection Dataset

Dataset Description

Source: UCI Machine Learning Repository – Occupancy Detection Dataset https://archive.ics.uci.edu/dataset/357/occupancy+detection

Features:

- Temperature (°C)
- Humidity (%)
- Light (lux)
- CO_2 (ppm)
- Humidity Ratio

Target (for evaluation only):

- Occupancy (0 = Empty, 1 = Occupied)

Problem Statement

In smart homes, automation depends on detecting human presence using sensor data. When labelled data is unavailable, unsupervised learning techniques can be used to cluster sensor patterns and infer occupancy to control appliances such as lights and fans.

Theory

Unsupervised learning algorithms group similar sensor readings into clusters without using labels.

1. K-Means Clustering

Partitions data into K clusters by minimizing intra-cluster distance.

2. Hierarchical Clustering.

Clusters are later mapped to occupancy labels using majority voting.

Algorithm

1. Load IoT dataset
2. Select sensor features
3. Normalize data
4. Split into training and testing sets
5. Train clustering models
6. Map clusters to occupancy labels
7. Predict occupancy
8. Apply automation rules
9. Evaluate models
10. Visualize results

Automation Rules

Condition	Action
Occupied & Light < 200 lux	Turn ON lights
Occupied & Temp > 28°C	Turn ON fan
Else	Turn OFF devices

Observations

- **K-Means shows better stability than hierarchical methods**
- **Agglomerative clustering is sensitive to scaling**
- **Hierarchical clustering provides useful structural insight**
- **Unsupervised methods perform lower than supervised models**

Result

Successfully implemented smart home automation using unsupervised clustering techniques and evaluated their performance using accuracy metrics, confusion matrice

Learning Outcomes

1. **Understand IoT sensor data characteristics** and how environmental parameters such as temperature, humidity, light, and CO_2 affect smart home automation.
2. **Apply unsupervised machine learning algorithms** (K-Means, Agglomerative, and Hierarchical clustering) to real-world IoT datasets.
3. **Perform data preprocessing techniques** including feature selection, normalization, and train-test splitting for sensor data analysis.
4. **Map clusters to meaningful real-world labels** (occupancy detection) using majority voting.
5. **Design rule-based automation logic** based on AI model outputs for controlling home appliances.
6. **Evaluate model performance** using accuracy metrics and confusion matrices.
7. **Interpret clustering behaviour visually** using dendrograms and comparative performance plots.