Explain how a Hidden Markov Model (HMM) could be used in your project by addressing the following:

Describe the Observations: What measurable data would the model use?
Type of HMM Problem: If you don't know the hidden states in advance, what kind of HMM task is this?
Training Algorithm:
a. What values are known at the start?
b. What values are unknown and need to be learned?
c. Parameter Updates: Which HMM parameters will your training algorithm update?

Example of Dataset
Smart Bee Colony Monitor

Unsupervised Anomaly Detection in Beehives

Observations: Measurable Data

Environmental Parameters

Temperature: Inside and outside hive (°C)

Humidity: For bee survival (%) - (most of the bees die during the rainy season due to humidity.

Sound Analysis: Bee activity patterns, queen presence, swarming indicators

Movement Detection: Bee traffic in/out of hive

Weight Sensors: Honey production tracking

Weather Data: Rainfall, wind, atmospheric pressure,(external humidity, atmospheric pressure, wind speed, and precipitation levels.

Also, the model will include temporal observations such as time of day, season, and day of the week to capture cyclical patterns in bee behavior.

Might be included(At the moment still doing further research on them)

Behavioral Parameters

Number of bees leaving and returning to the beehive

Hive Vibration Patterns: Colony health indicators

Sound Frequency Analysis: Different frequencies indicate different colony states

Entrance Activity: Traffic patterns throughout the day

HMMs Model summary Integration

Observations (O)

Measurable Data: (As I indicated their purpose above)

Temperature readings (continuous)

Humidity levels (continuous)

Sound amplitude and frequency (continuous)

Movement counts (discrete)

Weight measurements (continuous)

Weather conditions (categorical)

Processed Features:

Sound pattern classifications

Activity level categories (low, medium, high)

Environmental stress indicators

Productivity metrics

Hidden States (H)

Colony Health States**:**

Healthy & Productive: Optimal conditions, high activity

Stressed: Environmental pressure, reduced activity

Critical: Disease, queen loss, swarming preparation

Declining: Poor conditions, low survival probability

Recovery: Improving from stressed state

HMM Problem Type

Hidden states are not known in advance, this is a Parameter Learning task using the Baum-Welch Algorithm (Expectation-Maximization).

The actual health states of bee colonies are not directly observable and must be inferred from sensor readings, this constitutes an unsupervised learning problem specifically categorized as a parameter estimation task.

The problem type is most likely to be a Baum-Welch algorithm implementation where the model must simultaneously learn the sequence of the hidden state and the model parameters from observation sequences. This involves estimating the most likely sequence of hidden states given the observed sensor data, making it a combination of the decoding problem where finding the most probable state sequence and the learning problem where estimating model parameters..

Training Algorithm (assumption)

Values that are known from the start:

Observation sequences: Historical sensor data

Number of hidden states: 5 (assuming I will be using the one I mentioned above)

Observation space: (Still doing research on this)

Unknown Values to Learn:

Transition probabilities (A): P(state_t+1 | state_t) =This is supposed to define the likelihood of moving from one hidden state to another, is completely unknown and must be estimated from the temporal patterns in the data.

Emission probabilities (B): P(observation | state) =This will be representing the likelihood of observing specific sensor readings given a particular hidden state, requires learning from the correlation between observations and inferred states

Initial state probabilities (pie): P(initial_state) = This will indicate the likelihood of starting in each hidden state, will be likely to be estimated from the beginning of observation sequences

Parameter Updates

(Still researching)