

Using HMMs in a Crop Recommendation System

Project Title:

Intelligent Crop Recommendation Based on Environmental and Soil Dynamics Using Hidden Markov Models

1. Observations (Measurable Data):

The observations used in the HMM would be **daily or seasonal measurements** of the following features:

- Soil nutrient levels (e.g., Nitrogen, Phosphorus, Potassium)
- Temperature and humidity
- Rainfall
- pH value
- Light exposure
- Historical crop yields

These are the **visible, measurable data points** collected over time.

2. Type of HMM Problem:

Since the **optimal crop choice or soil state is not directly observable** and must be inferred from environmental conditions and past yields, this is a **Learning Problem** (also called the **Unsupervised HMM problem**).

We don't know the hidden states (e.g., "soil fertility level", "climate suitability state"), and must **learn both the model structure and parameters from the data**.

3. Training Algorithm:

a. Known Values at the Start:

- Sequences of observations (e.g., a year's worth of temperature, soil nutrients, etc.)

- Number of hidden states (assumed or chosen experimentally, e.g., “low fertility”, “medium fertility”, “high fertility”)

b. Unknown Values to Learn:

- Transition probabilities between hidden states (e.g., how likely soil fertility changes over time)
- Emission probabilities (likelihood of observing a certain temperature, rainfall, etc., in a hidden state)
- Initial state probabilities

These are learned using the **Baum-Welch algorithm**, a form of Expectation-Maximization (EM).

4. Parameter Updates:

The training algorithm updates the following HMM parameters:

- **Transition Matrix (A):** Probability of transitioning from one hidden state to another (e.g., from “low fertility” to “medium fertility”)
 - **Emission Matrix (B):** Probability of observing a set of environmental features given a hidden state
 - **Initial State Distribution (π):** Probability of the system starting in each hidden state
-

Conclusion:

In this project, an HMM can help **model the temporal evolution of environmental and soil conditions**, infer **hidden soil/climate states**, and use that to **predict or recommend the most suitable crops** for a given period. This approach is useful in regions where agricultural data is incomplete or noisy but still follows underlying patterns over time.