Uses of Hidden Markov Model in the my Community Yield Forecasting Network Capstone Project

Mission Statement

My mission is to empower Rwandan farmers with a smart, data-driven system that accurately predicts crop yields using real-time data and advanced machine learning techniques. By integrating valuable feedback from local farmers, I aim to create a continuously improving tool that enhances agricultural decision-making, boosts productivity, and supports sustainable farming practices across Rwanda.

Description of Observations

In my project, the HMM[1] would use measurable data collected from Rwandan farmers via the web app, including:

- **Yield Data**: Monthly or seasonal crop yield in kg/ha (e.g., maize, beans).
- **Environmental Variables**: Rainfall (mm), temperature (°C), and soil moisture (%) from local weather stations or farmer inputs.
- **Farmer Actions**: Fertilizer application (kg/ha) and planting dates, logged anonymously. These observations will serve as the visible outputs, reflecting the underlying agricultural states (e.g., crop health stages) that the HMM aims to model.

Type of HMM Problem

Since the hidden states such as crop growth phases (e.g., germination, vegetative, reproductive) or yield potential categories (e.g., low, medium, high) are not known in advance due to the diverse and evolving nature of farmer data, this is an **unsupervised learning problem** within the HMM framework. Specifically, it aligns with the **Baum-Welch algorithm** task, where the goal is to infer both the hidden states and the model parameters from the observed data without prior state labels.

Training Algorithm

a. **Values Known at the Start**:

- **Observation Sequence**: The sequence of measurable data (yield, rainfall, etc.) collected from farmers over time.
- **Initial Model Parameters**: A random initialization of the transition matrix (probabilities of moving between states), emission matrix (probabilities of observations given states), and initial state distribution, as starting points for the Baum-Welch algorithm.

b. **Values Unknown and Need to Be Learned**:

- **Hidden States**: The underlying crop health or yield potential states (e.g., growth phases or productivity levels) that drive the observed data.

- **Transition Probabilities**: The likelihood of transitioning between hidden states (e.g., from vegetative to reproductive phase) based on environmental changes.
- **Emission Probabilities**: The probability of observing specific yield or weather data given a hidden state.

Parameter Updates

The training algorithm, Baum-Welch, will update the following HMM parameters iteratively:

- **Transition Matrix**: Adjusted to reflect the true probabilities of state changes based on temporal patterns in the data (e.g., how often a crop moves from a healthy to a stressed state).
- **Emission Matrix**: Updated to better map observed variables (e.g., high rainfall to a specific yield range) to the inferred hidden states.
- **Initial State Distribution**: Refined to match the most likely starting state for a given crop season, improving the model's starting point for predictions.

Application in the Project

The HMM would enhance the crowdsourced model by identifying hidden agricultural states from farmer inputs, improving yield predictions over time. As more data is anonymously contributed, the model refines its parameters, enabling personalized forecasts and supporting the gamified engagement by rewarding farmers whose data improves prediction accuracy.

REFERENCE

1. G. Avinash, V. Ramasubramanian, M. Ray, R. K. Paul, S. Godara, G. H. H. Nayak, R. R. Kumar, B. Manjunatha, S. Dahiya, and M. A. Iquebal, "Hidden Markov guided Deep Learning models for forecasting highly volatile agricultural commodity prices," *Applied Soft Computing*, vol. 148, no. 111557, Jun. 2024.available:https://www.sciencedirect.com/science/article/abs/pii/S1568494624003314