

Capstone Integration: Hidden Markov Model in Angaza

Context

How a hidden Markov Model can be applied within Angaza, a machine learning driven system for reporting, monitoring, and tracking domestic violence and femicide in Kenya.

1. Observations

What measurable data would the model use? In the context of victim reports received via the Angaza portal, the HMM would use:

- Time-stamped incident descriptions.
- Sentiment scores and urgency levels.
- Number and types of uploads
- Repeated access or edits to previously submitted reports.
- Interaction history with the chatbot.
- External data, like court updates.

2. Type of HMM problem

Since in this case I don't have ground truth for the hidden emotional and safety states of the victim, the model will need to learn the hidden states from patterns in the observed data.

3. Training Algorithm

Known Values:

- Sequence of observable variables, i.e, user behavior over time
- Vocabulary and sentiment/urgency score mappings from pre-trained NLP models

Unknown values to learn:

- Transition probabilities between hidden states, e.g, from "at risk to "in danger".
- Emission probabilities: likelihood of observing a certain text/emotion/upload pattern given a hidden state.
- Initial state distribution (e.g, probability that a new report starts in "at risk" state)

I will use the Baum-Welch Algorithm to train the model.

4. Parameters to update

- Transition matrix (A): Probability of moving from one hidden state to another
- Emission matrix (B): Probability of an observable event given a hidden state.
- Initial state distribution (π): Probability of starting in each hidden state.

These updates will allow the system to deduce the hidden progression of victim states and trigger appropriate actions or alerts.