Capstone Project Title: Multilingual Bone Fracture X-ray Report Generation and Speech System for Low-Resource Communities

This capstone project will be in two parts, which is the X-RAY report generation and the multilingual speech system.
For the X-ray report generation, the observations are sequences of features extracted from key anatomical regions like the femur, tibia, spine, hip, or shoulder. These regions are identified using object detection models such as YOLO or Faster R-CNN. Feature vectors capturing local details like texture and edges are then extracted with a CNN and arranged in a consistent order. This sequence of feature vectors serves as input to the HMM, which uses them to infer hidden clinical states such as **"no fracture," "minor fracture," or "severe fracture."**
In the multilingual speech system, the observations are acoustic features like MFCCs and spectrogram frames extracted from short segments of recorded speech. These feature vectors form a sequence that captures how speech sounds change over time. The HMM uses this sequence to model transitions between speech units such as phonemes, enabling clear and natural spoken medical reports in local languages like Kusaal or Twi.

This is an unsupervised learning problem where the HMM must learn hidden states,such as the clinical labels or phonemes directly from observed data. The goal is to estimate the most likely hidden state sequences and model parameters (transition, emission, and initial probabilities).This is done using the **Baum-Welch algorithm,** allowing the system to uncover meaningful patterns without the need for labeled hidden states.

At the beginning of training, the only **known values are the observation sequences**, which are the Image feature vectors extracted from different anatomical regions (e.g., femur, tibia, spine) in the X-ray images and the Acoustic feature vectors, such as MFCCs or spectrogram frames, extracted from recorded speech.
The unknown values that the training algorithm must learn from the data are; **Transition probabilities**, **Emission probabilities and Initial state distribution.**

During training, the algorithm updates the key HMM parameters to better model the observed data. These include the transition probabilities, which represent the likelihood of moving from one hidden state to another; the emission probabilities, which capture how likely each observed feature (such as image or acoustic data) is generated by a specific hidden state; and the initial state distribution, which reflects the probability of starting in each hidden state. By iteratively adjusting these parameters using the **Baum-Welch algorithm,** the model improves its fit to the sequences of observations and uncovers the underlying hidden state patterns.