1. What are the vt , ht in this case? What is their domain?
2. What are the parameters we need to learn?
3. Where will we get training data from?
4. What is the inference step? Can it be done quickly?
5. Is the HMM a good model for this problem, what are possible extensions and what are the problems that may come with them?