

BTech(ICT) Semester IV: Probability and Random Processes (MAT202)

Special Assignment Abstract

Date: February 14, 2019 (Thursday)

- Area- Biology: Disease Progression Modeling Using Hidden Markov Model
- Group Members:
 1. Mudra Suthar (1741022)
 2. Hardi Trivedi (1741092)
 3. Rutvi Tilala (1741091)
- Background:
 - Diseases are the unavoidable part of our lives. There are diseases like common cold, which have minor and avoidable symptoms, But diseases such as AIDS and Alzheimers are unseen and unpredictable in nature.
 - It makes a lot of us wonder how some people get caught in the grip of these dreadful diseases and some escape. People always desire to know whether they are at greater risk of getting such diseases so that they can take precautionary measures.
 - Over the past many years, mankind has tried to find and understand this uncertain nature of these diseases by using mathematical tools like probability and statistics and thus tried to prevent the spread of these diseases.
- Importance of Topic:
 - People are always in the fear of how serious the diseases can become and at what rate. People are unaware of how can any disease take such a serious form and how fast does it propagate. This depends on various factors such as hereditary problems, infections, radiations, viruses, food habits etc and generally end up in not getting the disease diagnosed and an effective treatment.
 - Getting the detailed information about the diseases and their progression in the patients can help to take precautionary measures.
- Inference:
 - What exactly you are going to do?
 - We are planning to use Hidden Markov Model in such a way that it effectively uncover underlying statistical patterns in disease progression by defining different disease states. The connections between the states are to be governed by a set of transitional probabilities. Hidden Markov Models consists a set of the states which are statistical having associated probability distributions called the observation probability density functions. These observations are typically a multidimensional vector consisting of a set of features called the HMM feature vector. The observation density functions can either be defined as continuous or discrete.
 - We will likely use continuous distributions based on Gaussian Mixtures for our project. We will define n clinical stages that will lead us to the transitional probability from state to state.
 - Let A_{ij} be the transitional probability from state i to state j , $b_i(x)$ be the observation probability density function of state i , x be the HMM feature vector, and C_i be the a-priori probability of starting in state i . Since disease progresses with time, we propose a left-to-right topology, where we will consider each HMM state as a disease stage. A transition that occurs from one state to the

next indicates disease progression while a transition to the same state indicates no progression. We will also include a transition to a previous state i . Depending on the specific disease that we are modeling, such a transition can indicate disease regression due to, for instance, the presence of a potential treatment or because the disease regressed naturally.

- References:

1. S.R. Eddy, Hidden Markov Models and Large-Scale Genome Analysis, Trans. American Crystallographic Assoc., 1997.
2. Rafid Sukkar, Senior Member, IEEE, Elyse Katz, Yanwei Zhang, David Raunig, and Bradley T. Wyman