CS 3600 Project 3 Wrapper

CS3600 - Spring 2022

Due April 10th 2022 at 11:59pm EST via Gradescope

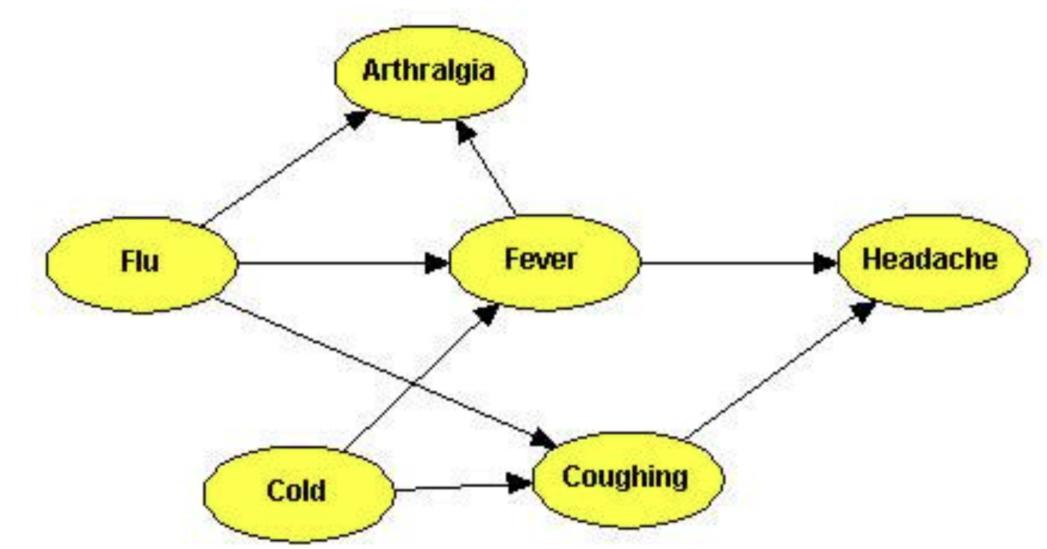


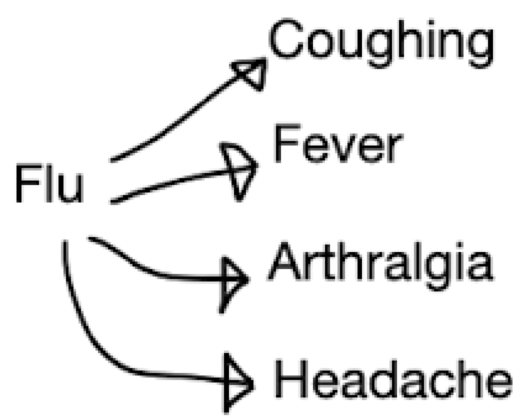
Figure 1: Example Bayesian network for medical diagnosis.
Source: http://song.bayesian.net/index.php/Bayesian_net

Probabilistic inference over Bayesian networks is a standard AI technique for medical diagnosis. Bayesian networks represent complex causal relationships between patient information, medical conditions, and symptoms. Probabilistic inference allows us to compute diagnostic queries, determining the likelihood of medical conditions given observed symptoms as evidence.Use the example Bayes net above as a prompt for the following questions.

Question 1

Recall that the naïve Bayes assumption is that no effects of a cause are also causes of each other. If two effects are correlated it is because they are related to the same, underlying cause. The naïve Bayes model provides an alternative representation for diagnostic inference. Draw a Bayes net representing the naïve Bayes model for diagnosing *Flu* given its symptoms (assume the symptoms of *Flu* are every successor of *Flu* in the Bayes net in Figure 1). Which model (the Bayes net in Figure 1 or the naïve Bayes model that you've constructed) is a richer representation? That is to say, is there anything we can represent with one model that we cannot represent with the other model?

Answer:



The bayesian net in figure one is richer in this case. We can see that the Bayesian net in figure one can show the correlation between symptoms like headache, fever and coughing but since the Naive Bayes assumption assume every symptom is independent, it does not show all relationships. The Naive Bayes assumption might misinterpret some of the probabilistic results. However, the Naive Bayes assumption of conditional independence allows for more efficient computation. It uses less computation power.

Question 2

$SICK_{t-1}$	$P(SICK_t = T \vert SICK_{t-1})$	$P(SICK_t = F \vert SICK_{t-1})$
T	0.7	0.3
F	0.5	0.5

Table 1: Transition Probabilities

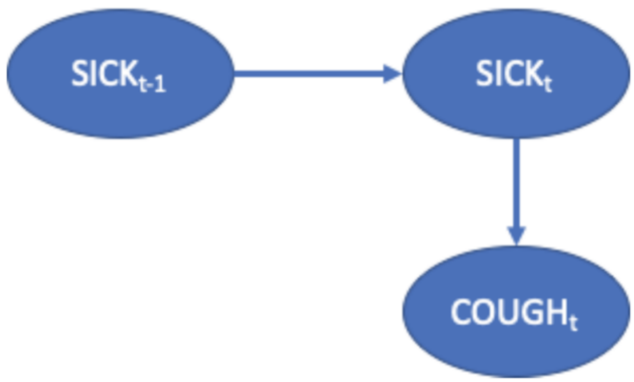


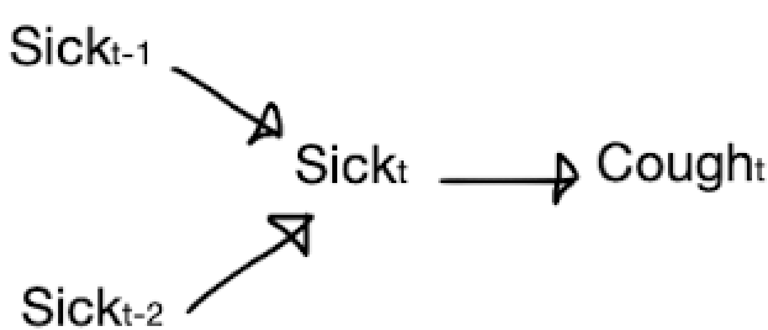
Figure 2: First Order Markov Dynamic Bayes Net

The traditional Dynamic Bayes Net has an unobservable random variable X_t that has a single parent of the value of X_{t-1}, which is the value of X at the previous time step. For example, SICK$_t$ is conditioned on SICK$_{t-1}$. This can capture a relationship such as "when one is sick, the probability is high that one is still sick at the next time step, and when one is not sick, one can become sick or stay well with equal probability". See the image for an example. However, if one were to use this Bayes network to predict the future, the model may conclude that people become sick randomly and then stay sick. This setup does not account for second-order effects, such as: "after one is sick for a while, the probability is high that one stops being sick". A 2-Markov assumption states that an unobservable random variable X_t is conditioned on X_{t-1} and X_{t-2}. **Using a time step equal to a week, draw a 2-Markov Dynamic Bayes Network that captures the intuition that one can become sick at any time**. When one is sick one is likely to remain sick unless they have been sick for two weeks, at which time they are likely to cease being sick. When one is sick, the probability of cough is high and when one is not sick, the probability of cough is low. **Show all the conditional probability tables; make up reasonable numbers to express the relationships described above.**

Answer:

$Sick_t$	$P(Cough_t = T \mid Sick_t)$	$P(Cough_t = F \mid Sick_t)$
T	0.9	0.1
F	0.1	0.9

$Sick_{t-1}$	$Sick_{t-2}$	$P(Sick_t = T \mid Sick_{t-1}, Sick_{t-2})$	$P(Sick_t = F \mid Sick_{t-1}, Sick_{t-2})$
T	T	0.1	0.9
F	T	0.5	0.5
T	F	0.7	0.3
F	F	0.5	0.5



Question 3

Medical diagnosis with Bayesian networks are currently used as a *decision support systems* by healthcare professionals. An expert can input patient information and observed symptoms, and the decision support system outputs a set of possible diagnoses with associated likelihoods, but the final diagnosis decision is up to the medical professional. Why should we require a human supervisor to accept or override the decision of the AI diagnosis system? Name two (2) potential sources of error or unaccounted for situations for these Bayes net diagnosis models that are mitigated by having a trained healthcare professional make the final diagnosis decision.

Answer: It seems to me that the first reason is that Bayes-net-based medical diagnoses are not good at analysing high dimension data. So if the symptom of a disease is too complicated, the network might not do well in terms of the accuracy of the result. The second reason is that the model could be outdated. Since new research and paper get published every day, the level of correlation between symptoms and disease could change. Bayes-net does not update itself and requires refactoring to keep updated. Therefore, having a capable and professional doctor to override is always a good idea. The first potential error comes from the data. The information regarding the symptom is input by a human. And if the input is incorrect, the result can be incorrect as well. The second potential error happens when the disease is brand new to both the model and the trained healthcare professional. Since the model is basically the mathematical model that builds using existing knowledge, it does not respond to unseen diseases. Therefore, do not give an accurate result.

Question 4

Publicly accessible online services often use databases and symptom matching to inform users of possible medical conditions given a list of symptoms. These services do not provide diagnosis likelihoods. Could providing a free online service with Bayes-net-based medical diagnosis have negative impacts on human behavior? Could they have positive impacts? If you answered yes to either question, give one example. If you answered no, explain why not.

Answer: I think that online services that inform users of possible medical conditions given a list of symptoms could induce negative human behaviour. It is because the service might not provide accurate results and might delay patients getting help. Since users might not give an accurate description of their symptoms, the result in Bayes-net-based diagnosis can be faulty, garbage-in-garbage-out. The second reason is that Bayes-net-based medical diagnoses are not good at analysing high dimension data. It means that if some medical condition's symptoms are too complicated, the network does not do well in providing the accurate result. On the other hand, they surely also have some positive impact. Firstly, Bayes-net reasoning is easy to trace back for future analysis. It is a model that represents a mathematical model graphically. So it is possible for doctors to review the reasoning behind this diagnosis and proceed forward. Secondly, this service does well in analysing low dimension data, which means that it could tell diseases, which have strong symptoms accurately and fast.