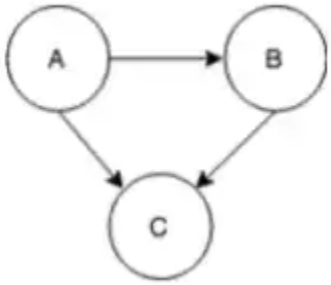
Bayesian probabilistic models

A Bayesian network is a probabilistic graphical model which uses directed acyclic graph (DAG) and a set of parameters to infer outcomes of events. The DAG is composed of nodes and edges. The nodes represent the random variables, X = X1, . . . Xi, . . . Xn from the domain, which is the universal set of events. Their conditional dependencies are reflected in the edges, for example, Xi -> Xj represents the direct dependencies between the two variables. The structure of the network captures qualitative relationships between nodes. In particular, two nodes should be connected directly if one affects or causes the other, with the edge indicating the direction of the effect. For example, there are three nodes, A, B, C which represents three different events. Suppose event A and event B could cause event C happening and event A has a direct effect on event B, then these three nodes could form a Bayesian network with node A pointing to B and C, and node B pointing to C.



The joint probability function for the above example is equation 1. Generally, the joint probability function is formed as equation 2. The P a(Xi) represents the set of parents of variable Xi in the DAG. Acyclicity of the DAG guarantees the product in equation 2 is a coherent probability distribution.

Key Concepts:

Bayesian Probability:

Bayesian probability is a framework for modeling uncertainty by incorporating prior beliefs (prior probabilities) and updating them with observed data to obtain posterior probabilities.

In social computing, Bayesian models are used to represent and infer probabilistic relationships between variables.

Probabilistic Models:

Bayesian probabilistic models describe probabilistic relationships between variables in a social context. They provide a principled way to model uncertainty and make predictions.

Common Bayesian Probabilistic Models:

1. Naive Bayes Classifier:

Used in text classification and sentiment analysis, it assumes that features are conditionally independent given the class label.

2. Latent Dirichlet Allocation (LDA):

A topic modeling technique that uncovers hidden topics in a collection of documents. It has applications in content recommendation and understanding user interests.

3. Bayesian Networks:

Represent probabilistic relationships between variables using directed acyclic graphs. Bayesian networks are employed in various social computing tasks, including collaborative filtering and social network analysis.

4. Hidden Markov Models (HMMs):

Used to model sequences of events with hidden states. In social computing, HMMs are applied to understand user behavior patterns, such as clickstream analysis.

5. Bayesian Matrix Factorization:

Applied in collaborative filtering to factorize user-item interaction matrices. This approach is effective for recommendation systems.