CAUSAL NETWORKS

Causal network illustrates the causal dependencies of all the components in the network A causal relationship exists when one variable in a data set has a direct influence on another variable. Thus one event triggers the occurrence of another event. A causal relationship is also referred to as cause and effect o The ability to identify truly causal relationships is fundamental to developing impactful interventions in medicine policy, business and other domains.

o Often in the absence of randomized control trials, there is a need for causal interne purely from observational data

 One possible explanation for correlation between variables where neither causes the other is the presence of confounding variables that influence both the target and a driver of that target. Unobserved confounding variables are severe threats when doing causal inference on observational data.

o A causal generalization, eg, that smoking causes lung cancer is not about an particular smoker but states a special relationship exists between the property of smoking and the property of getting lung cancer.

As a causal statement, this says more than that there is a correlation between the two properties o Some causal conditions are necessary conditions:

The presence of oxygen is a necessary condition for combustion; in the absence of oxygen there is no combustion.”Cause” is often used in this sense when the elimination of the cause is sought to eliminate the effect (what's causing the pain?)

1.2 Structural Causal Models (SCMs)

Structural causal models represent causal dependencies using graphical models that provide an intuitive visualisation by representing variables as nodes and relationships between variables as edges in a graph.

- SCMs serve as a comprehensive framework unifying graphical models, structural equations, and counterfactual and interventional logic.

- Graphical models serve as a language for structuring and visualizing knowledge about the world and can incorporate both data-driven and human inputs.
- Counterfactuals enable the articulation of something there is a desire to know, and structural equations serve to tie the two together.
- SCMs had a transformative impact on multiple data-intensive disciplines (e.g. epidemiology, economics, etc.), enabling the codification of the existing knowledge in diagrammatic and algebraic forms and consequently leveraging data to estimate the answers to interventional and counterfactual questions.

2. Bayesian Networks (BNs)

2.1 Directed Acyclic Graph (DAG)

- A graph is a collection of nodes and edges, where the nodes are some objects, and edges between them represent some connection between these objects.
- A directed graph, is a graph in which each edge is orientated from one node to another node. In a directed graph, an edge goes from a parent node to a child node.
- A path in a directed graph is a sequence of edges such that the ending node of each edge is the starting node of the next edge in the sequence.
- A cycle is a path in which the starting node of its first edge equals the ending node of its last edge. A directed acyclic graph is a directed graph that has no cycles.

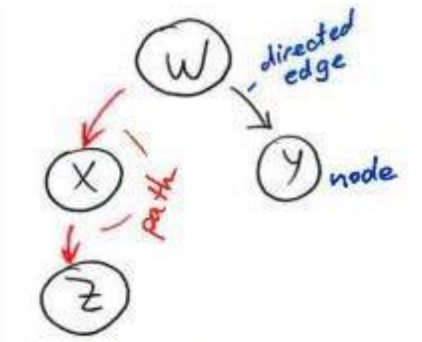

Figure 1: A simple directed acyclic graph.

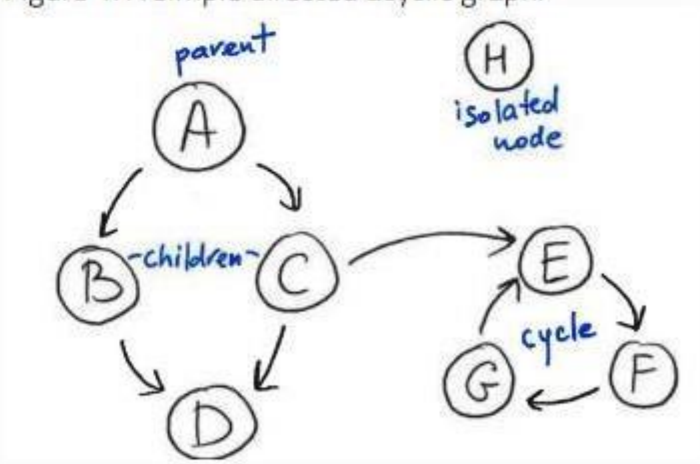

Figure 2: A more complex graph with a cycle and an isolated node. This graph can be turned into a DAG by removing one of the edges forming a cycle: (F, G), (E, F) or (G, E).

2.2 What Bayesian Networks are and are not

What are Bayesian Networks?

Bayesian Networks are probabilistic graphical models that represent the dependency structure of a set of variables and their joint distribution efficiently in a factorised way.

Bayesian Network consists of a DAG, a causal graph where nodes represents random variables and edges represent the the relationship between them, and a conditional probability distribution (CPDs) associated with each of the random variables.

If a random variable has parents in the BN then the CPD represents

$P(\text{variable|parents})$ i.e. the probability of that variable given its parents. In the case, when the random variable has no parents it simply represents $P(\text{variable})$ i.e. the probability of that variable.

Even though we are interested in the joint distribution of the variables in the graph, Bayes' rule requires us to only specify the conditional distributions of each variable given its parents.

The links between variables in Bayesian Networks encode dependency but not necessarily causality. In this package we are mostly interested in the case where Bayesian Networks are causal. Hence, the edge between nodes should be seen as a cause -> effect relationship.

Let's consider an example of a simple Bayesian network shown in figure below. It shows how the actions of customer relationship managers (emails sent and meetings held) affect the bank's income.

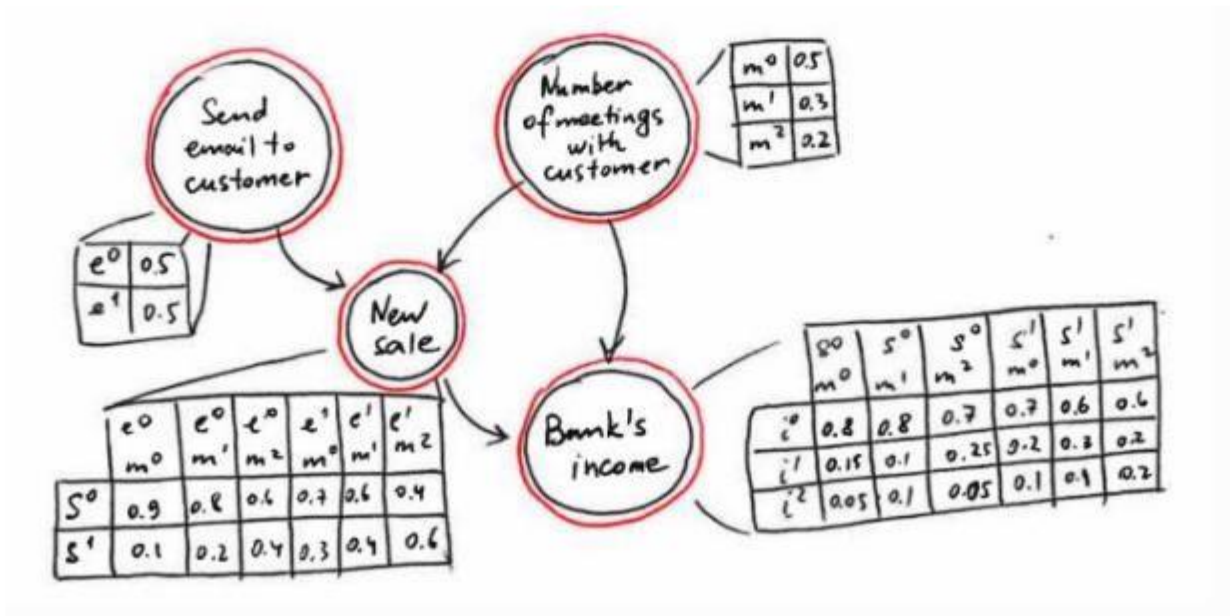

Figure 3: A Bayesian Network describing a banking case study. Tables attributed to the nodes show the CPDs of the corresponding variables given their parents (if present).

New sales and the number of meetings with a customer directly affect the bank's income.

However, these two drivers are not independent but the number of meetings also influences whether a new sale takes place.

In addition, system prompts indirectly influence the bank's income through the generation of new sales. This example shows that BNs are able to capture complex relationships between variables, represent dependencies between drivers, and include drivers that do not affect the target directly.

Steps for working with a Bayesian Network

BN models are built in a multi-step process before they can be used for analysis.

1. **Structure Learning**. The structure of a network describing the relationships between variables can be learned from data, or built from expert knowledge.

2. **Structure Review**. Each relationship should be validated, so that it can be asserted to be causal. This may involve flipping / removing / adding learned edges, or confirming expert knowledge from trusted literature or empirical beliefs.

3. **Likelihood Estimation**. The conditional probability distribution of each variable given its parents can be learned from data.

4. **Prediction & Inference**. The given structure and likelihoods can be used to make predictions, or perform observational and counterfactual inference. CausalNex supports structure learning from continuous data, and expert opinion. CausalNex supports likelihood estimation and prediction/inference from discrete data. A Discretiser class is provided to help discretising continuous data in a meaningful way.

What can we use Bayesian Networks for?

The probabilities of variables in Bayesian Networks update as observations are added to the model. This is useful for inference or counterfactuals, and for predictive analytics. Metrics can help us understand the strength of relationships between variables.

- The sensitivity of nodes to changes in observations of other events can be used to assess what changes could lead to what effects;
- The active trail of a target node identifies which other variables have any effect on the target.

2.3 Advantages of Bayesian Networks

- Bayesian Networks offer a graphical representation that is reasonably interpretable and easily
- explainable;
- Relationships captured between variables in a Bayesian Network are more complex yet hopefully more informative than a conventional model;
- Models can reflect both statistically significant information (learned from the data) and domain expertise simultaneously;