Bayesian Networks for Key Driver Analysis

Motivation

The director of marketing of a major airline just received the latest customer satisfaction report from their marketing research agency. They surveyed 2,000 travellers and asked about their traveling experience. The data was analysed with Key Driver Analysis and they found that flight attendant and gate personnel friendliness, snacks & drinks offered, and priority boarding all had significant impact on customers' satisfaction. The director wondered though: Would customers see flight attendant friendliness in a more positive way if the flight attendants served snacks? Or, if attendants were friendlier, would they not mind so much the absence of snacks? Standard Key Driver Analysis (e.g., linear regression, Shapley Value, etc.) can't answer these questions. Bayesian Net analysis possibly can.

Introduction

Identifying the key drivers that impact the success of your products and business is a vital practice in marketing research. By identifying what, for example, drives metrics such as overall satisfaction, perceived value and how likely your customers are to recommend your product, to their friends, family, and colleagues, allows your business to identify which improvements in your product or service will yield the biggest gain in these metrics. Marketing decision makers can then subsequently decide where to invest in first.

There are, in our opinion, two important reasons to consider Bayesian Networks:

1. One, although there are many models which can be used to run driver analyses, such as linear and logistic regression, decision trees, etc. (e.g., see Vriens, Vidden & Bosch, 2021). Most don't provide a deeper understanding of how independent variables might be influencing each other. In classical regression modelling, for example, we can simulate how a change to an independent variable will impact the dependent variable, but we can't see the effect that it will have on the other independent variables. This can be a substantial limitation because it may under or over-estimate the importance of a driver.
2. Two, many models may provide biased estimates when multicollinearity is present in the data. This weakness can be, in large part, avoided when using Bayesian Networks.

By using Bayesian Networks for Key Driver Analysis, we can identify the causal impact variables have on one another and visualize these relationships. In this whitepaper, we'll provide a simple explanation of what Bayesian networks are, provide a simple example of using a Bayesian network for key driver analysis, and outline the benefits and drawbacks of using a Bayesian network.

What are Bayesian Networks?

A Bayesian Network (BN) is a graphical model consisting of nodes, each of which representing a variable, and directed edges (arrows). If there is an edge from Node A pointing to Node B, for example, then this indicates that some part of the distribution of the variable represented by Node B is influenced by (or conditionally dependent on) the distribution of the variable represented by Node A. Although this sounds complicated, it much easier to interpret in an example.

Take a simple scenario, for example, where you have some data from your customers about their thoughts on your products design quality, efficiency, and their overall satisfaction with the product. A BN may model the interactions between these variables in the following structure seen in Figure 1 below:

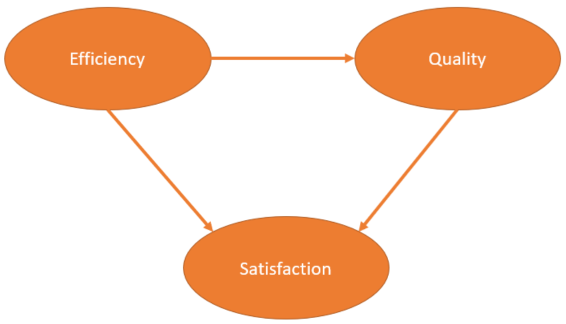
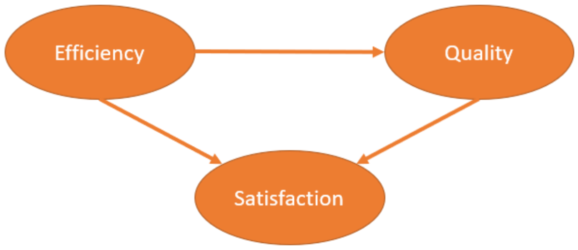


Figure 1. *A simple illustration of a Bayesian Network.*

From this, we can see that efficiency has a direct impact on quality and customer satisfaction, whereas quality only has an impact on customer satisfaction. What we can also infer from this structure is that some of the influence the customer's opinion on the quality has on their overall satisfaction can be explained through the search efficiency. In other words, this indicates that efficiency is a driver (not necessarily a very important driver) for quality. Both product efficiency and product design quality are drivers for customer satisfaction.

Efficiency=high	Efficiency=low
0.80	0.20

Efficiency	Quality=high	Quality=low
low	0.20	0.80
high	0.80	0.20



Efficiency	Quality	Satisfaction=high	Satisfaction=low
low	low	0.00	1.00
low	high	0.30	0.70
high	low	0.75	0.25
high	high	0.95	0.05

Figure 2. *Example of conditional probabilities.*

We can visualize the impact of each variable on every other variable by looking at their conditional probability tables. For efficiency, we see that 80% of customers deem the efficiency of the product as “high”. We can see that for quality, if we know that the customer deems the efficiency as “high”, then there is again an 80% chance that they will deem the quality as “high”. Finally, we can see that if both efficiency and quality are scored highly, then satisfaction has a 95% chance of being considered “high”.

This is just a simple example of Bayesian Networks, but hopefully it clarifies the idea behind them. The key importance is that you able to uncover all interactions between variables and the indirect effects.

Naturally, in the real world you wouldn’t have a pre-defined structure already. This is often found by automatically trying many different possible structures and saving the one which fits the data the best. This gets a little complicated when defining a search strategy and a metric to score the model, see Tsamardinos, Brown & Aliferis (2006) and Scanagatta, Salmerón & Stella (2019) for further discussion.

Using Bayesian Networks for Driver Analysis

From Figure 2 above, we can learn a lot about the distribution of our variables and how they interact with one another, but it does not provide us with simple, quantitative metrics which show the impact they have on one another. One approach to attain these values is by using the *Average Treatment Effect* (ATE). The ATE tells us how much a target variable is impacted when changing (intercepting) another variable. Intuitively, it can answer the question of “If I increase efficiency from low to high, how much of an impact will this have on customer satisfaction?”. When running ATE with the above model, we find that the impact efficiency has on customer satisfaction is approximately 0.85. Comparing this to design quality on customer satisfaction, we only see an impact of 0.21. We can see this visualized in the table 1 below:

Intervention Variable	Target Variable	ATE
Efficiency	Quality	0.6
Efficiency	Satisfaction	0.85
Quality	Satisfaction	0.21

Table 1. *An example with average treatment effects.*

Benefits and drawbacks of Bayesian Networks

Bayesian networks have numerous advantages over traditional regression-based driver analyses:

1. BN can be used to identify richer causal relationships between independent & dependent variables
2. As a result, it identified both direct and indirect effects.
3. Multicollinearity is not an issue
4. The entire data-generating mechanism is uncovered: We learn the structure of the drivers and their interactions
5. If there is knowledge about the data structure, then this can be integrated into the Bayesian network

However, there are some drawbacks, namely:

1. Computational complexity makes it difficult to find the true best structure
2. Requires discretization of continuous variables
    a. A simple survey response in [1, 2, 3, 4, 5] can be easily modelled by a discrete distribution
    b. However, very granular scores (e.g., 3.57) will lose some information
3. A Bayesian Network will struggle learning the structure of data with many variables. It is generally best applied on less than 25-30 variables
4. Inference may be weaker than other simple models, such as Logistic Regression and decision trees

References and further reading

1. Vriens, M., Vidden, C. & Bosch, N. (2021). The benefits of Shapley-value in Key-driver analysis. In: Applied Marketing Analytics, 6, 3, 269-278.

2. Tsamardinos, I., Brown, L. E., & Aliferis, C. F. (2006). The max-min hill-climbing Bayesian network structure learning algorithm. Machine learning, 65(1), 31-78.

3. Scanagatta, M., Salmerón, A., & Stella, F. (2019). A survey on Bayesian network structure learning from data. Progress in Artificial Intelligence, 8(4), 425-439.

For more information contact:

Chad.vidden@teamkwantum or marco.vriens@teamkwantum.com