# Bayesian Networks for Marketing Analytics: Understanding Customer Behavior Through Probabilistic Modeling

**Christian Joseph Fiore**
Master's Degree in Artificial Intelligence, University of Bologna
christian.fiore@studio.unibo.it

March 23, 2025

**Abstract**

This mini-project demonstrates the use of Bayesian Networks within marketing analytics in order to learn from the behavior of consumers and predict responses to marketing campaigns. I use probabilistic modeling strategies to discover dependences among customer attributes and purchase behavior. My comparison of the use of exact inference methods over sampling-based methodologies showed that even with the network non-polytree structure, exact inference provided a better result than sampling-based methods.

## Introduction

### Domain

Marketing analysis involves modeling customer behavior and predicting their reactions to marketing activities. Traditional methods depend to a great extent on segmentation and classification techniques, which cannot capture the probabilistic nature inherent in customer decision-making processes. This study applies Bayesian Networks (BNs) to illustrate causal dependencies between customer attributes and reactions to marketing efforts. BNs are particularly valuable in this domain because they explicitly represent conditional dependencies among variables, thereby enhancing the intuitive understanding of complicated relationships.

### Aim

This project aims to apply a Bayesian Network in marketing analytics to demonstrate its real-world application in a data-intensive field. My particular goals are:

- Demonstrating how Bayesian Networks can successfully represent consumer behavior in marketing, a data-driven field by nature Using probabilistic inference to predict customer outcomes in practical marketing situations.
- Comparing the efficiency of exact inference with sampling-based inference techniques to identify the most suitable technique for practical applications

### Method

I employed the pgmpy library to model my Bayesian Network and run probabilistic queries. The process involved the steps below:

- Preprocessing a marketing analytics dataset to properly deal with categorical variables, discretize continuous variables, and improve memory efficiency.
- Development of a Bayesian Network structure from expert knowledge in consumer behavior.
- Training the model with Maximum Likelihood Estimation of the available data
- Running queries through the application of two inference methods: Variable Elimination (exact) and Weighted Sampling (approximate).
- Analyzing and comparing results to make marketing inferences

### Results

The Bayesian Network was very effective for marketing analysis in that it identified statistically significant patterns in consumer behavior with fewer data compared to conventional machine learning algorithms. Despite simplifications to the datasets due to memory limitations, the model still retained analytical rigor while its inference speeds were remarkable. Exact inference methods surpassed sampling-based approaches despite the network's complex non-polytree structure, indicating the method's computational efficiency at dealing with small marketing datasets.

## Model

My Bayesian Network is a non-polytree model of consumer behavior in marketing contexts. It is complicated by the fact that it has multiple paths between nodes through purchase and campaign response variables—a structure that is typically challenging for exact inference.

The network includes four categories of nodes:

**Customer Attributes:** Demographics (Income, Age) discretized into 4 bins via quantile binning to have equal distribution.

**Engagement Metrics:** Customer_Days (tenure length), Recency (days since last purchase), Complain (whether customer complained).

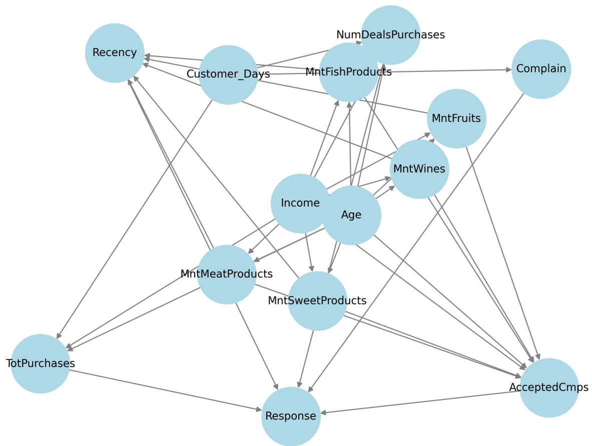


Figure 1: Bayesian network structure modeling customer attributes and purchasing behavior

**Purchase Behavior:** TotPurchases (total purchases), NumDealsPurchases (deals purchases), and various spending categories (Mnt* nodes).

**Campaign Outcomes:** AcceptedCmps (0-5, which past campaigns are accepted), Response (whether customer has responded to current campaign).

The model captures domain knowledge in that Income/Age drive buying behavior and campaign acceptance. Purchase behaviors drive response to the campaign, and recency and complaints influence the likelihood of acceptance.

CPTs were learned using Maximum Likelihood Estimation. To deal with memory constraints, I cut down the dataset by:

- Omitting demographic variables that have little predictive importance
- Adding purchase columns to TotPurchases
- Consolidating campaign acceptance markers into a single AcceptedCmps variable

## Analysis

### Experimental setup

I ran seven different probabilistic queries to understand consumer behavior and marketing efficiency:

1. Influence of income level on campaign acceptance
2. Purchase behavior in relation to age and income
3. Probability of complaints by customer tenure
4. Deal buying and response rate interaction
5. Recency and response probability correlation
6. The joint distribution of campaign acceptance and purchases given income
7. The joint distribution of response and complaints based on age and deals

Every question was run using both exact inference (Variable Elimination) and approximate inference (Weighted Sampling with 10,000 samples). I compared the accuracy of the resulting probability distributions and the computational cost of both methods.

### Results

My comparison of inference techniques resulted in a surprising result with important implications for probabilistic modeling in marketing applications. In spite of the intricate non-polytree structure of my Bayesian Network, exact inference (Variable Elimination) consistently outperformed approximate inference in computational cost.

Due to constrained computational power, I encountered problems with the initial training of the model with the original *data_cleaned.csv* dataset that called for a systematic approach in reducing dimensionality and simplifying the dataset. This preprocessed substantially reduced the complexity of conditional probability tables and thereby allowed the Variable Elimination algorithm to achieve greater accuracy and perform better than Weighted Sampling for both query types.

These findings defy conventional wisdom about inference method selection, proposing that for small-domain real-world problems, highly optimized exact inference is the method of choice even under modest computational resources.

## Conclusion

This project demonstrates the strength of Bayesian Networks in probabilistic relational modeling for real world application such as marketing analytics. This approach has obvious benefits over subsymbolic AI techniques for smaller data sets and manageable network complexity.

The comparison of inference techniques yielded an unexpected outcome: even with the network's non-polytree structure, exact inference consistently performed better than sampling-based approaches in terms of efficiency and accuracy. The implication is that for small-domain problems, in which nodes have a limited number of parent nodes, optimally designed exact inference may be better than approximation techniques.

Bayesian Networks are especially helpful when there is limited data availability since they give interpretable probabilistic conclusions without the high data demands of most machine learning methods. The graphical structure enables easy reasoning about conditional dependencies while ensuring computational tractability.

## Links to external resources

Dataset source: `https://www.kaggle.com/datasets/jackdaoud/marketing-data`

## References

Koller, D., and Friedman, N. 2009. *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA: MIT Press.

(Koller and Friedman 2009)