

# Common algorithms in marketing: Decision trees

**Zylo**tech™

Customer Data & Analytics Blog

## Common algorithms in marketing: Decision trees

Janet Wagner, on July 18, 2019 | 3 minute read



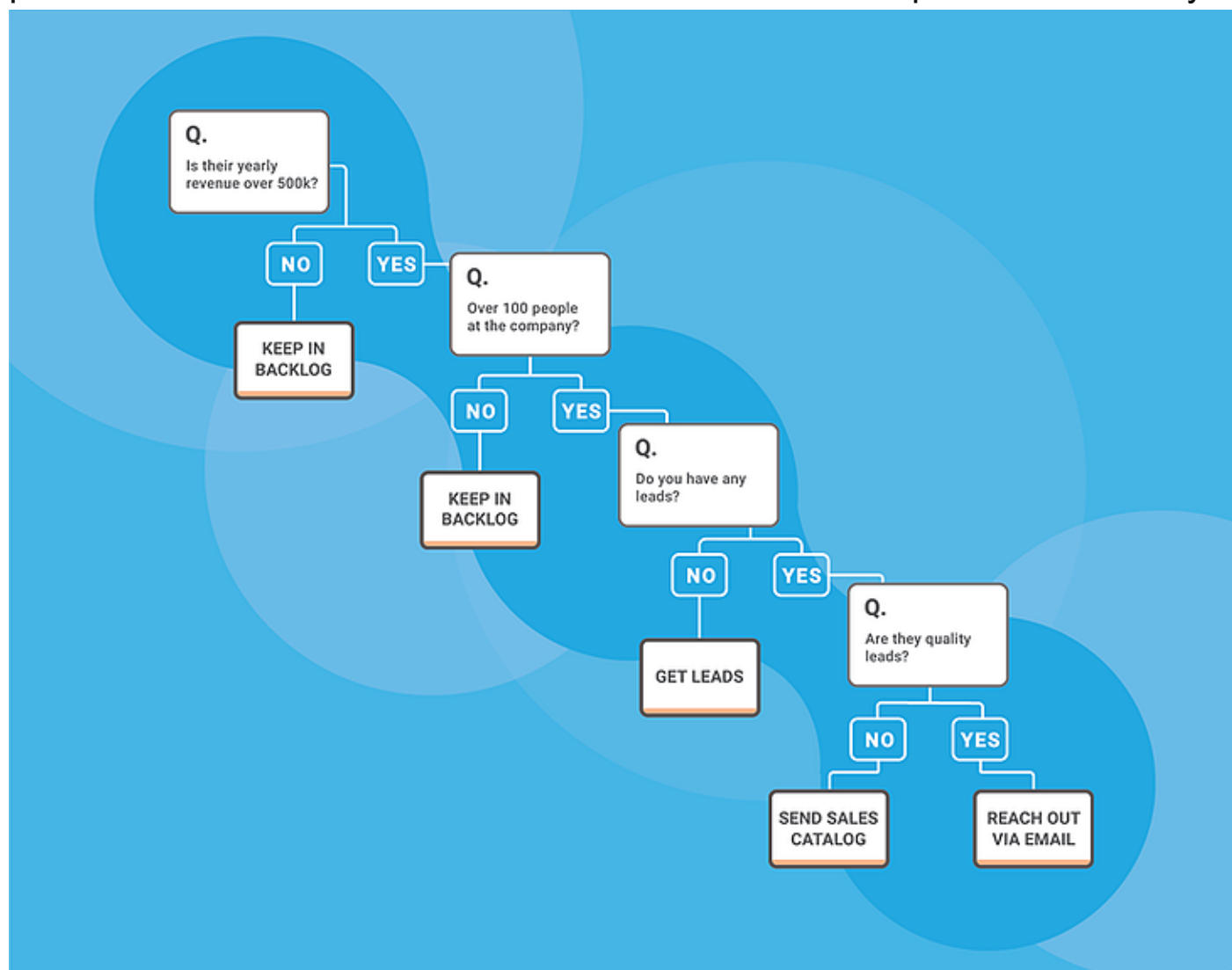
Organizations that have incorporated machine learning (ML) into their platforms (or plan to) are finding that there are many decisions to be made when it comes to models and algorithms. One of those decisions is which

algorithms should be used for which applications. Decision tree algorithms are commonly used in marketing. This post highlights some of the advantages and disadvantages of decision trees. This post also includes a few examples of how decision tree algorithms are used in marketing.

## What are decision trees?

Decision tree algorithms are used primarily for [predictive modeling](#). Decision trees consist of a series of checks that branch into different potential outcomes. Trees can branch based on whether a field has some value, numeric thresholds, or when a specific condition has been met. Decision trees can be used for classification and regression- they generate rules that are used to predict outcome variables which can be either continuous (for regression trees) or categorical (for classification trees). The

process of decision trees forms a tree structure when represented visually.



## Advantages of decision trees

- Decision trees are good at handling numerical variables and categories of variables such as true/false and match/no match.
- They are [nonparametric](#) ML algorithms, which means there is potentially an infinite number of parameters. Also, the algorithms do not make assumptions about data distribution so they can learn any functional form from the training data set.
- Tree-based models notice relationships. If you build a decision tree model for selling a house, a tree will notice that some of the variables are related. For example, the model would notice that if a house has a

higher number of bedrooms, it will have a larger total of square feet. Some models, like regression models, see each variable as independent so the relationships would not be recognized.

- Decision trees are interpretable (to a point – see disadvantages), allowing findings to be communicated in a manner that non-technical audiences can understand.

## Disadvantages of decision trees

- Decision trees tend to have problems with [overfitting](#), especially when a decision tree is extremely deep. There are a number of techniques to combat overfitting, such as setting a max depth, [pruning](#), and [bagging](#).
- Decision trees require data points that are [linearly separable](#).
- Slight changes in the training dataset can produce unexpected outcomes from the model. This is because each child node depends on a parent node- an entire tree structure can change if the first split feature is changed.
- Decision trees are interpretable [up to a point](#). The deeper a tree, the more nodes there are. And the more nodes there are, the more difficult it will be to convert the tree into something interpretable.

## Marketing use cases

Decision trees are used in marketing for a wide range of use cases, two examples:

### Improve outbound marketing efforts

Decision trees can be used to analyze customer data and answer marketing questions such as “which outbound marketing activities should we do more of?” Using decision trees, marketers could predict which customers are

most likely to respond favorably when receiving a promotional email or a



sales catalog in the mail.

## Increase customer loyalty

Decision trees can be used to determine which customers are most likely to spend more money at a business if they are given a loyalty rewards card. The model could generate a target value that predicts the probability of each customer spending more with the card. For example, a value of “1” would mean the customer is *likely* to spend more and a value of “0” would mean the customer is *unlikely* to spend more.

## Open source ML libraries

There are a lot of open source ML libraries available, some of which include decision trees. If you’re interested in experimenting (mostly in Python) with decision tree algorithms check out [scikit-learn](#), [TensorFlow](#) (Boosted Trees), and [Apache Spark MLlib](#).

Looking for more open source libraries? This [blog post](#) features open source libraries and tools data scientists in marketing should check out.

*Janet Wagner is a Zylotech contributing writer.*

*If you liked this post, check out our recent [blog post: Self-Learning AI enables intelligent recommendations](#).*