

Market Segmentation with Clustering

Introduction

In this lesson, we'll learn about one of the most popular use cases for clustering in the business world -- market segmentation!

Objectives

You will be able to:

- Explain market segmentation and how clustering can be used for it

What is Market Segmentation?

Perhaps the most common use case for clustering algorithms in the real world, **Market Segmentation** refers to using **Cluster Analysis** to segment a customer base into different *market segments* using the clustering techniques we've learned.

Consider the following scenario: You're a movie executive, and you have a new superhero film coming out. You need to decide how to best allocate your advertising budget in order to attract the most customers. This film is a sequel, so you have good demographic data on who went to see the last film. The advertising options available to you are TV, newspaper, radio, and internet. How do you best allocate your advertising budget to ensure that the movie does as well as possible?

The answer depends on your data. A regression analysis on last year's data can give you a general idea of how much you can expect to make overall, assuming that there aren't major differences between last year and this year. However, regression just tells you what you can expect *overall* -- what if we're trying to optimize where we spend our money, rather than just predict what the returns will be, based on the overall amount of money we spent?

The answer lies in knowing who your customer is. All forms of advertising are not consumed equally by every age group or demographic. By identifying **segments** in our customer data, we can look for trends that identify one group or another, and create personalized regression models for each group.

In order to understand this better, let's take a sample question that market segmentation can help us answer. For our TV advertising budget, we still have to decide what channel to run our commercials on. What effect will advertising on the Disney channel have on a person's likelihood of coming to see our superhero movie? If the person in question is 12 years old, then it's probably very likely that our

commercial convinces this person to see our movie. But what about if they're 68 years old? In that case, advertising during a cartoon on the Disney channel might not be the most effective way to reach that person. If we're worried about reaching this customer, the first question we should ask is what kind of customer they are. In the case of a superhero movie, we can likely assume that all things equal, a 12-year-old child is more likely to be interested in seeing a superhero movie after seeing our commercial than a 68-year-old, so we should probably pay attention to what the data tells us about how 12-year-olds are affected by each type of media advertisement we can use!

Two potential customers deep in conversation about what movie to see

You can bet that movie studio executives have complex, well-defined models to predict their Return on Investment (RoI) for things as granular as advertising on Disney, versus advertising on the History channel. This is because different market segments of customers behave differently, and market segmentation allows us to zoom in on those groups!

Identifying Market Segmentation

At the most basic level, market segmentation allows us to look at our data and identify which customers belong to which groups. Once we have this information, we can examine each individual segment and use it to answer important questions and build individual, targeted models for each segment.

Once we understand our market segments, then we can begin making informed decisions that are specific to each segment. So how do we find these market segments?

With clustering, of course! By definition, market segments are groups within our dataset with substantive differences between them. A segment only matters to us if it is different from other groups -- we don't really care about identifying the segment of children with red hair versus children with brown hair in our data if they both act the same way and have the same level of interest in our movie. Before data scientists became commonplace, this sort of segmentation was usually handled by marketers using their intuition about their customer base to create *customer personas*, and then seek out data to back up their assumptions. As data scientists, we know that the best option is not to seek data to confirm our beliefs -- instead, it is to pull our beliefs from evidence in the data. Clustering provides a great way for us to allow the data to tell us what is and isn't significant -- lest we get caught up chasing down market segments that aren't actually all that different -- or worse, don't actually exist at all!

Segmentation and Targeting

In modern business analytics, segmentation is only the first step.

After we've identified the different market segments, the next step is to build individualized strategies to **Target** them! In the movie example we used above, we would first start by answering questions such as "which market segment is most valuable to us?" This can be answered through research or through analyzing our data, or a combination of both. Once we realized that the 12-to-18-year-old demographic is most valuable to us, we can then decide how to target them in the most effective way possible. This brings us back to our earlier question -- how do we allocate our advertising budget? If we've used regression to determine that we're most likely to get the return on investment for our advertising dollars with the 12-to-18-year-old age group, then our next step is to determine which ad channels are most effective to us. We'll likely find that TV advertisements and internet ads are very effective at reaching this particular market segment, but radio is less effective (since a solid portion of the target segment can't yet drive), and newspaper ads are unlikely to reach them at all (because when is the last time you saw a 12-year-old read a newspaper?).

The third step in this process is a bit outside the scope of clustering. This is where the marketing team really shines -- figuring out how to position our product to make it both as desirable as possible to a given segment, while also making our product stand out from competitors.

Let's look at one more example to consider what this looks like: car advertisements!

Scenario: You are the newest data scientist at Tesla Motors. Next year, you are introducing a new SUV in the \$30-50k price range. The SUV is roomy, spacious, fast, and affordable, in addition to having a very high safety rating and a ton of technological bells and whistles. One day, Elon Musk asks you (presumably, on Twitter) who your valuable market segments are, and what parts of the car he should highlight in several upcoming interviews. How do you answer this question?

Presumably, the first thing you would do is to look at the results of your market segmentation and identify the most profitable market segments to target. Once you know who these segments are, you can target them with ads -- but this only brings us to the second step in our diagram above. The third step means personalizing these ads to have maximum impact on a given targeted segment. Is your target market middle-class families? Then maybe it makes sense to highlight the car's affordability, space, and safety rating. What about if your target is upper-middle-class customers between 30 and 40 years of age that enjoy luxury cars? In that case, you'd probably focus on the speed, luxury, and looks of the car, because they're more likely to care about these qualities than the others.

When you know your market segment, you can market to them in the most effective way possible!

Step 3 in this process is usually done with the help of survey data, under the umbrella of *User Research*. This is not something that data scientists typically have to worry about too much, as it is a

different domain of expertise. However, the first two stages are very much something that data scientists can expect to do multiple times in their career!

Summary

In this lesson, we learned about how cluster analysis can be applied to determine market segmentation, and how these market segments are used in the real world to plan and execute effective business strategies!