

Drill:

Let's say we work at a credit card company and we're trying to figure out if people are going to pay their bills on time. We have everyone's purchases, split into four main categories: groceries, dining out, utilities, and entertainment.

What are some ways you might use KNN to create this model? What aspects of KNN would be useful? Write up your thoughts in submit a link below.

Given that KNN focuses on similarities that are shared among neighbors, it would make sense to look at “neighbors” who default on their bills and have similar spending habits on groceries, dining out, utilities, and entertainment.

KNN will find similar clusters of behavior to be able to figure out personas from a data set. For example, are there similar spending habits with regards to discretionary spending (dining out and entertainment) vs non-discretionary spending (groceries and utilities).

You could also look at similarities in clusters of behavioral spending defaulters vs. non-defaulters in non-discretionary spending. For example, how do defaulters compare to non-defaulters in mean spending for groceries and utilities?

Another way to use KNN would be to look at spending over a billing cycle to see if defaulters vs. non-defaulters tend to spend in cyclical patterns that might predict default.