Week 8 Homework –Power Company Case Study

Question 18.1

Describe analytics models and data that could be used to make good recommendations to the power company.

Here are some questions to consider:

- The bottom-line question is which shutoffs should be done each month, given the capacity constraints. One consideration is that some of the capacity – the workers' time – is taken up by travel, so maybe the shutoffs can be scheduled in a way that increases the number of them that can be done.
- Not every shutoff is equal. Some shutoffs shouldn't be done at all, because if the power is left on, those people are likely to pay the bill eventually. How can you identify which shutoffs should or shouldn't be done? And among the ones to shut off, how should they be prioritized?

Think about the problem and your approach. Then talk about it with other learners and share and combine your ideas. And then, put your approaches up on the discussion forum, and give feedback and suggestions to each other.

You can use the {given, use, to} format to guide the discussions: Given {data}, use {model} to {result}.

Have fun! Taking a real problem and thinking through the modeling and data process to build a good solution framework, is my favorite part of analytics.

Problem Definition:

Some customers of a power company pay bills, some don't. Those who don't pay bills either genuinely cannot pay the bills (and there are programs to support such customers), and then there are those who can pay, but just don't pay. For the latter type of non-paying customers, the power company wants to shut off their power supply. But there are logistical issues with shutting the power – Shutoffs have to be done manually and workers have to go to each location manually to shut down power; and Time taken to travel to each location is more load than the company can handle. So, there is a need to develop an analytics-based approach to tackle this problem. I am going to approach the problem using Given {Data}, Use {Model} to {Result} type solution.

All the available data?

1. Credit Score,
2. Number of years with the company,
3. Payment History,
4. Estimated Annual Income of the household,
5. Whether there has been any previous shutoffs/debt collection related to the account,
6. Total amount owed to the power company,
7. Average energy consumption each month,

8. Peak energy consumption, and day of the month of peak energy consumed,
9. Average, peak (and day of peak consumption) energy consumption of the neighborhood,
10. Total number of defaulted accounts in a ‘n’-mile radius from the defaulted account,
11. Distance from the power company to the defaulted household.

I’m going to split the problem into 3 subsets:

1. Classify the type of customer (will pay, can’t pay and won’t pay);
2. Rank the risk of shutting off vs. keeping power on;
3. Build an optimization model to minimize the number of personnel to send for shutting off power and time taken, distance travelled from one place to another.

I. <u>Classify the type of customer (will pay, can’t pay and won’t pay):</u>

Given:

1. Credit Score,
2. Number of years with the company,
3. Payment History,
4. Estimated Annual Income of the household,
5. Whether there has been any previous shutoffs/debt collection related to the account

To: Classify the type of customer (will pay, can’t pay and won’t pay)

Using:

1. Classification (SVM): This will help classify customers based on their ability and willingness to pay the bills on time. Some additional factors that can be considered are whether the customers classified as non-paying customers default only in certain months/seasonality of their default payments.
2. Logistic Regression: This will help determine the probability of a customer to pay the bill a little late with a small penalty. We can perhaps use a Ridge/Elastic Net to determine the most important factors in the regression model.

II. <u>Classify the risk of shutting off vs. keeping power on:</u>

Given:

1. Total amount owed to the Power company,
2. Annual household income,
3. Payment History,
4. Average energy consumption each month,
5. Bill history for each month,
6. Distance from the power company to the defaulted household.

To: Rank the risk of shutting off vs. keeping power on

Using:

1. Linear Regression: Again, using Ridge/Elastic Net, determine the most important factors for the regression model. Weight the model with highest weights from the previous month's bill and energy usage to lowest to the earliest available bill month and its energy usage history.
2. Exponential Smoothing: This will help forecast the next month's energy consumption. This will help determine the risk of keeping power on vs. shutting off the power.

III. Optimization (and Simulation):

Given:

1. Peak energy consumption, and day of the month of peak energy consumed,
2. Average, peak (and day of peak consumption) energy consumption of the neighborhood,
3. Total number of defaulted accounts in a 'n'-mile radius from the defaulted account,
4. Distance from the power company to the defaulted household,
5. Amount of available personnel and salary of each personnel.

To: Build an optimization model to minimize the number of personnel to send for shutting off power and time taken, distance travelled from one place to another.

Use:

1. Clustering (k-means): Using k-means clustering, we are able to cluster households where power has to be shut off based on available number of technicians, their salary.
2. Optimization: In such a problem, the objective function of the optimization problem would be to minimize the distance to the household per personnel (thereby minimizing cost per personnel). Constraints would be number of available personnel, number and distance of houses from the company location, and time to shut off power (and perhaps even reinstall power after the bill is paid). Using the output from the clustering problem will help with the optimization problem.
3. Simulation: We can use ARENA to build a simulation model to determine the number of personnel needed using probability distributions for average time taken per personnel from one household to another