Homework 14: ISYE 6501 - Introduction to Analytics Modeling

Question 20.1

Prompt

Describe analytics models that could be used to help the company monetize their data: How could the company use these data sets to generate value, and what analytics models might they need to do it?

There are lots of good answers, and I want you to think about two types – at least one of your answers should be based on just one data set, the one they've collected internally on customer browsing patterns on the web site; and at least one of your other answers should be based on combining more than one of the data sets.

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}".

Here are the three data sets to consider:

DATA SET #1 (purchased from an alumni magazine publisher) - first name - last name - college or university attended - year of graduation - major or majors - marital status - number of children - current city - email domain - financial net worth - binary variables (one for each interest in the publisher's long list of various sports, activities, hobbies, games, etc.) showing whether each one was or wasn't listed by each person

DATA SET #2 (purchased from a credit bureau) - first name - middle name - last name - marital status - sex - year of birth - current city - whether they ever owned real estate - email domain - list of monthly payment status over the last five years for credit cards, mortgages, rent, utility bills, etc. – for each month and each payment: - what type of payment it was – for credit cards, it would say "Visa", "American express", etc., not just "credit card" - how much was owed - how much was paid - whether the person was considered to be in default

DATA SET #3 (collected by the company using web site tracking code) - title - first name - middle initial - last name - credit card type - credit card number - list of products purchased in the past, with date of purchase and ship-to address - which web pages the person looked at - how long the person spent on each page - what the person clicked on each page - estimate of how long the user's eyes spent on each page viewed (for customers where the software was able to take over the device's camera)

Solution

The optimal solution would be a combination of data sets. Assuming you can combine data set #1 and #2,

Considering Data Set #1, and Data Set #2, you can combine the demographic and interest data with the financial data. The best use would be a propensity scoring model with logistic regression or consideration of additional machine learning algorthms like using random forest trees, or XGBoost. By combining data sets you can predict high-value customer segments and target them with premium offerings or marketing campaigns. To implement the data into the model, merge data sets on common identifiers (e.g., name, email domain, marital status). You could also have additional features such as disposable income (derived from financial net worth and payment behavior), hobbies, and creditworthiness. Next, train a propensity scoring model to predict the likelihood of customers purchasing premium products or services. After, you can classify or segment the customer base into tiers: high-propensity (priority targets), medium-propensity, and low-propensity customers. This allows for determining the value generation. You can design targeted advertising campaigns for high-value segments, reducing customer acquisition costs. Depending on the interests, then you can also develop premium product bundles tailored to interests and financial capacity. This also allows to strengthen relationships with high-value customers through personalized loyalty programs, further increasing lifetime value.