ISYE 6501, Week 13 HW

Question 1

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

Response –

The problem statement is to identify what data is needed, which models will be used and what results are to be analyzed for a power company to identify a potential customer who may default a payment, decide what course of action to execute for a nonpaying customer based on their expected power usage and how to effectively run the shut downs maximizing the utilization of their resources.

1. Data

The data required for assessing a customer's ability to pay is –

- Credit score
- Income
- Duration of customer's account
- Bill payment history
- No. of defaulted payments
- Duration of default (if the customer pays back in a couple of days)
- Address of the customer
- Power usage history
- Time required for shutdowns
- Number of power company staff
- Office locations of power company (Starting point of staff who perform shutdowns)

2. Models

Multiple analytical models are required for this problem. The important aspects of the problem that are worth modeling are listed below.

- Potential non-paying customers
- Expected power usage
- Decision to shutdown
- Executing the shutdowns

2.1. Potential non-paying customers

Potential non-paying customers is either a classification problem or a clustering problem depending on the quality of the data available. Using unsupervised learning techniques like k-means clustering can provide quick results but the quality of results may not be accurate. Using a classification technique, like SVM or KNN, can provide a better segregation of non-paying customers from the entire customer base.

Date input for this model

- Credit score
- Income
- Duration of customer's account
- Bill payment history
- No. of defaulted payments

Output of the model –

Whether the customer pays the bill or may not pay the bill

2.2. Expected power usage

The amount of power that a potential non-paying customer might use can be modeled with time-series model if the customer has considerable history with the power company. However, newer customers' power usage might be difficult to model due to lack of history. A hybrid approach can be used to estimate the usage of a customer. Based on a threshold for the usage history, customers can be again classified into customers with enough usage history and not enough usage history. For customers with enough usage history, time series model can be applied. For customers with insufficient usage history, factor based regression model can be used as all the factors for modeling may not be available.

Data inputs

- Output of model from Section 2.1 above which identifies the non-paying customers
- Power usage history

Model Output

Expected power usage of the customer which gives the cost of leaving the power on.

2.3. Decision to shut down

The starting point of decision to shut down will be the cost of leaving the power on compared to the cost of shutdown. However, the cost of shutdown is not constant for each customer. For customers in the same vicinity location, the cost of shutdown will be lower compared to customers who are spaced out by longer distances. Even though customers are not far away, constraints like power company staff work hours, the duration it takes to shut down the power based on other constraints like weather (clear skies vs rainy day), access to the shutdown location etc also affect the cost of each shutdown.

Since travel times between shut down points can have significant effect on the decision for shut down, all the non-paying customers can be clustered based on cost of leaving power on (calculated from power usage model in 2.2) and the physical address of the customer. Then for each cluster, an optimization model could be run to decide a yes or no decision for the power utility workers to carry out the shutdowns. Simulation can also be used to account for the variability in the shutdown timings, travel time between shutdown locations etc.

The optimization model's objective function will be to maximize the savings (cost of power usage – cost of savings) for a customer whose decision is a "YES" for shutdown.

Data inputs for the model –

- Output of model from Section 2.2 above which estimates the cost of leaving the power on
- Address of the customer
- Number of power company staff
- Office locations of power company

Result

Decision to carry out a shutdown for a customer or not

2.4. Executing the shutdowns

The maximum expected value of savings for the company comes is driven by minimizing the cost of shutdowns. Cost of shutdown can be minimized by increasing the staff for shutdowns. The tradeoff here is the operational income increased by having more staff. Hence, a parametric study could be carried out using a simulation model to analyze how the savings for power company can change with changing staff numbers.

Data inputs for the model –

- Output of model from Section 2.3 above which provides decision for a shutdown
- Number of power company staff
- Office locations of power company

Result

Number of staff needed to maximize the savings for the power company.

A flow chart is shown in the following page to summarize the methodology explained above.

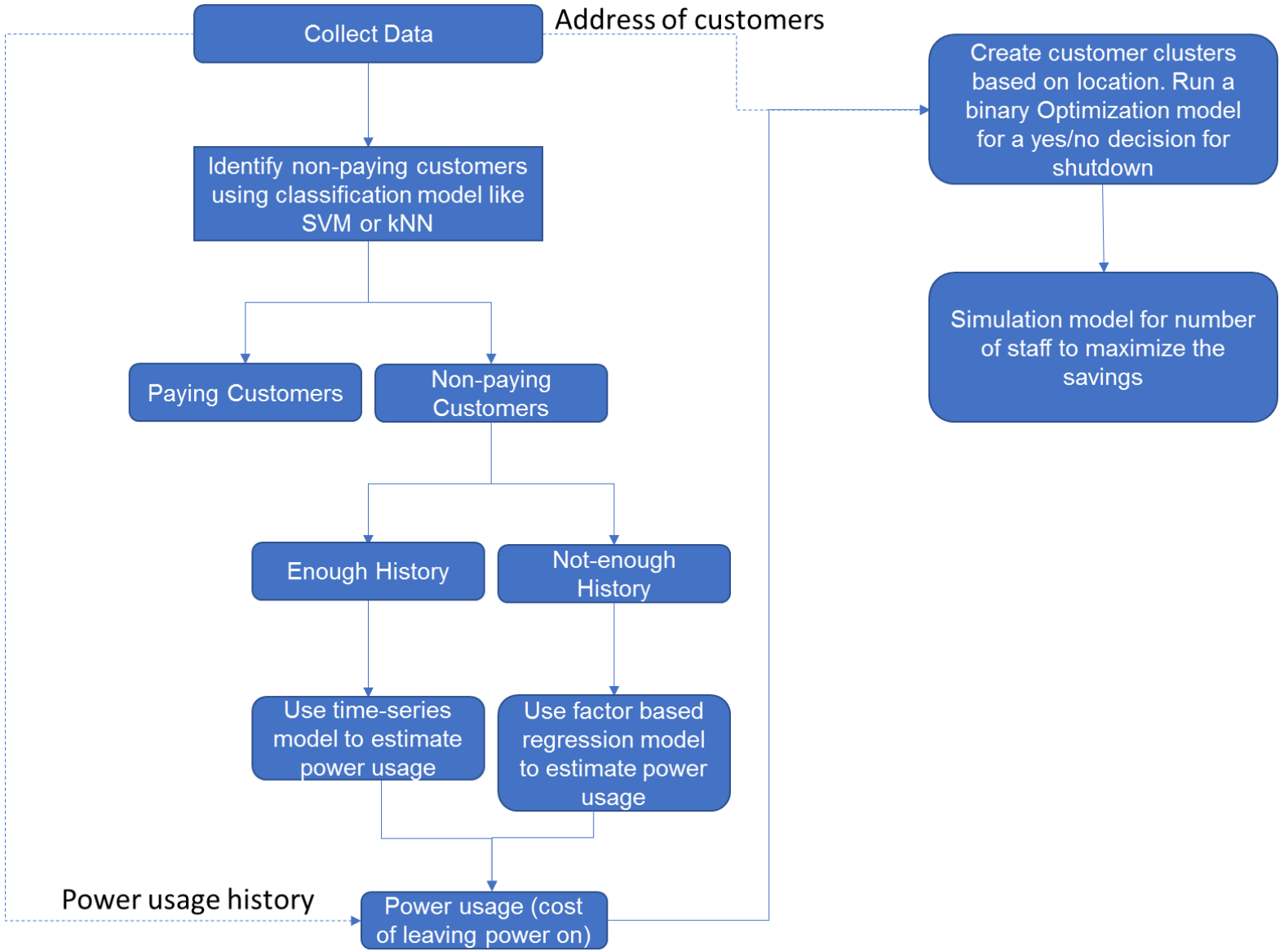

Figure 1 Flowchart for power company problem to identify non-paying customers and the decision for shutdowns