

Dear [AD]

Based on our team discussion, I've identified the overall structure and roadmap as to how we can help our client PowerCo go about navigating their customer churning issue. We will require different data sets to analyze various possibilities, but we will focus primarily on price sensitivity patterns.

We will use a binary classification model (supervised learning) to construct our predictive model to identify customers at risk of leaving PowerCo. Ideally, we require the following datasets:

- Demographics
  - o User type (non-commercial households, small businesses, large commercial buildings, etc)
  - o Location (urban or rural areas)
- Time dependant data
  - o Usage patterns
    - Historical patterns
    - Seasonal patterns
  - o Historical churn rate and timing
  - o Past discounts
    - Timing of providing discounts
- Competitor price information (if available) and timing of provider switch

The above information should be enough to create a relational database that gives us enough flexibility in investigation the specific reasons to customer price sensitivity. The relational database should give enough insight into relations between customer churn rate to various factors such as time dependency, location, usage pattern, competitor availability.

We plan to use binary classification which can conveniently be used at the beginning of every month to identify potential 20% discount targets. The model ideally takes account and accurately identifies the insights mentioned above. The efficacy of the model can then be evaluated using A/B testing methods. The model would be, at random, deployed for different customers. The A/B testing should provide insight into whether the outcome is favourable for PowerCo.

Best,  
Logan