Dear Sir/Madam,

Hypothesis: churn is driven by customers’ price sensitivity

Step 1. Utilise machine learning techniques to test variables impact on churn.

We can have the churn data as our target variable y, with 1 being stopping the service and 0 being stay with the service.

Then consider all factors that could affect churn rate and use them as the independent variables. So for each SME client, we will need the following data from them: location (country, city), current price, price difference compares to last month/year, gas and electricity monthly usage, total length they have been using PowerCo’s services, number of competitors, etc.

With these data we can build regression model and by calculating the weights for each variable we can see how significant these variables are affecting the target variable. The p-value of price-related variable can help us test the hypothesis.

Moreover, we can also build decision tree model to see what kind of variables leads to churn.

Step 2. Sensitivity test for the discounting strategy

To test the discounting strategy offered by our clients we can use sensitivity test.

First, we can split the clients into different groups randomly.

Second, assign different number of discounts to different control group.

Finally, compare the outcome from different control group to test the correlation between discount and recovery from churn.

These are my initial ideas about this project. Please let me know if other questions, thanks.

Kind regards,

Yuting