

Subject: PowerCo: Test whether churn is driven by price sensitivity.

Dear Iman Karimi,

To test client's hypothesis: **whether churn is driven by price sensitivity**.
We will need to predict customer's likelihood to churn and understand the effect prices have on churn rates.

Data needed:

- **Customer data:** characteristics of each client i.e. previous energy consumption, enterprise etc.
- **Price data:** historical prices the client charge to customers over time
- **Churn data:** whether customer has churned.

Workflow (iterative process):

1. **Data cleaning:** modifying/removing any incorrect, incomplete, irrelevant data.
2. **Data preprocessing:** normalization/standardization for better convergence, encoding categorical values, feature selection by whittling down predictors to a smaller set that is more informative.
3. **Build Binary classification model:** (e.g. Logistic regression, random forests, KNN, neural nets) pick the appropriate model after evaluating tradeoff between accuracy, complexity, and interpretability.
4. **Model validation:** Split data into training, validation, and test and score model performance.

From the model, we will understand the degree of impact prices have on churn rates. This will help us assess the effectiveness of the client's discounting strategy.

Context:

Client is a major gas and electricity utility.

- Supplies to corporate, SME, and residential customers.

Significant churn problem

- Drive by power-liberalization of the energy market in Europe

- Problem is largest in the SME segment

Client's Hypothesis

Churn is driven by price sensitivity.

Client wants to try a discounting strategy

- Offering customers at high propensity to churn a 20% discount

Regards,
Patrick Guo