# **Customer Churn Analysis**

### **Background information**

PowerCo is a major gas and electricity utility company that provides services to corporate, SME (Small & Medium Enterprise), and residential customers. The energy market in Europe has undergone power liberalization, which has resulted in significant customer turnover, particularly in the SME segment. It is reasonable to assume that changes in pricing have an impact on customer churn. Therefore, it would be valuable to identify customers who are more likely to churn at their current price, which can be achieved through the development of an effective predictive model.

In addition, for customers identified as being at risk of churning, offering a discount might serve as an incentive for them to remain with PowerCo. The head of the SME division is contemplating the implementation of a 20% discount, which is considered substantial enough to discourage almost any customer from churning, especially those who prioritize pricing. The client intends to utilize the predictive model on the first working day of each month to determine which customers should be offered the 20% discount. Formulate the hypothesis as a data science problem and lay out the major steps needed to test this hypothesis and the data that you would need from the client and the analytical models you would use to test such a hypothesis.

### Approach for problem :

Analyzed thoroughtly the client situation and formulated the hypothesis as a data science problem. Adding to this, I have identified the major steps need to test this hypothesis effectively. This will allow us to apply appropriate analytical models and methodologies for testing the hypothesis.

**Hypothesis :** The churn of SME customer is influenced by the price sensitivities. Specifically, Increase in pricing can significantly impact the likelihood of churn among SME customers.

1. Data Gathering:

To test this hypothesis, I would need the following data from the client:

1. Historical customer data: This should include information such as customer demographics, contract details, billing information, payment history, and churn status.
2. Pricing data: I require details on the pricing structure, including the timing and magnitude of price changes for each customer.
3. Data Exploration:

Performing exploratory analyses on relevant fields can provide valuable insights into customer churn behavior. Some key analyses to consider include:

1. Churn rates: Calculate churn rates for different customer segments (corporate, SME, residential) and analyze their distributions over time.
2. Pricing patterns: Explore historical pricing data to identify any patterns or trends in price changes. This can help us understand the impact of pricing adjustments on customer behavior.
3. Customer demographics: Investigate the demographic characteristics of churned customers compared to those who have remained loyal. This analysis can uncover any correlations between demographics and churn.
4. Data Preparation:

Create a structured dataset where each row represents an individual customer and each column contains relevant features and target variables. The dataset should include:

1. Customer attributes: Demographic information, contract type, tenure with PowerCo, etc.
2. Usage patterns: Energy consumption levels, billing frequency, etc.
3. Pricing variables: Price change history, price sensitivity indicators, etc.
4. Target variable: A binary indicator of churn (1 for churned, 0 for retained).
5. Feature Engineering:

Based on the available data, derive additional features that capture price sensitivities and other relevant factors influencing churn behavior. These features could include:

1. Price change indicators: Quantify the magnitude and frequency of price changes experienced by each customer.
2. Price sensitivity metrics: Develop metrics or indices that capture the price sensitivity of customers based on their historical behavior, such as responsiveness to price changes or historical churn patterns associated with price adjustments.
3. Automated Machine Learning (AutoML):

We can work on different model and select best model manually. Else, we can utilize AutoML techniques to automatically search for the best predictive model that estimates the likelihood of churn. AutoML platforms can evaluate a wide range of models, including Logistic Regression, Decision Trees/Random Forests, Gradient Boosting, and many others. The AutoML process will handle the model selection and hyperparameter tuning, optimizing the model's performance and accuracy. I can also go improve the best model further after evaluation.

1. Model Evaluation:

Split the dataset into training and validation sets to evaluate the performance of the predictive model. Use suitable evaluation metrics, such as accuracy, precision, recall, and F1 score, to measure the model's effectiveness in predicting churn.

1. Interpretation and Insights:

Analyze the results to gain insights into the relationship between price sensitivities and churn. This includes identifying significant features, assessing their impact on churn likelihood, and understanding the characteristics of customers who are more likely to churn due to pricing factors.

1. Discount Strategy:

Utilize the developed predictive model on the 1st working day of each month to identify customers at risk of churning. Offer the 20% discount selectively to those customers who are both predicted to churn and identified as price-sensitive, maximizing the impact of the discount while minimizing its overall cost.

1. Monitoring and Refinement:

Continuously monitor the outcomes of the discount strategy by tracking churn rates and customer responses to the discount offer. This will allow us to assess the

effectiveness of the 20% discount in incentivizing customers to stay. Additionally, I will regularly update the predictive model by incorporating new data and refining the features to improve its accuracy and predictive power over time.

In summary, my approach to test the hypothesis of churn being driven by customers' price sensitivities involves gathering historical customer and pricing data, performing exploratory analyses, preparing a structured dataset, engineering relevant features, building a predictive model using appropriate techniques, evaluating model performance, and analyzing results. I propose implementing a discount strategy based on the model's recommendations and continuously refining it. I kindly requested the client's assistance in providing the necessary data, including historical customer data and pricing information. By following this approach, I can gain insights, identify at-risk customers, and implement a targeted discount strategy.