Hi,

I understand to problem then we are trying to solve for PowerCo and I have a proposal as to how we can approach that problem.

The business problem is how we can identify factors that influence customers to churn so that we can predict which customers are most likely to churn. In the end, we are trying to minimize customer churn by taking actions such as customized discount offer to customers we identify as most likely to churn.

My approach is quite straightforward. First, we identify factors that may affect customer churn. We will do this through consultation with stakeholders and experts, and through our internal brainstorming sessions that include exploratory data analyses. Second, we gather datasets that capture such factors and process them. Third, we will build predictive model for customer churn. Lastly, we will build reports and dashboard to inform our findings and help stakeholders monitor the potential for customer churn.

Regarding hypothesis testing, we can compare data about customer's power usage data (or other data) from the month of which the churn data is available with the data from the previous month. The status on whether there are significant difference between the two months or not can be used as feature for our predictive modelling. And then such status can be used as a feature for our predictive model. Of course, we can also build predictive model without such hypothesis testing. Because maybe our model would be enough to capture the pattern of such chronological data.

Data that I propose we gather to be used as features include: the economy (interest rate, inflation, etc), government power subsidy, competitor's price. For the data regarding our customers, we can use data about the size of our SME customer's business, how long they have been our customers, records of their complaints, etc.

Those are all the points I have on the matter. Look forward to improve and finalize the approach further.

Many thanks,

Nugi