

## **Aily Case: EnergyInc**

*The following is a fictional case study. It will test your ability to handle data and perform statistical/machine learning analyses as well as your ability to communicate your findings and derive commercial insight from your technical work.*

*You may perform the analyses using any computational language you wish. Please submit your code along with your presentation and the requested results file by the date agreed with the recruitment team.*

*Please work through the below in your own time and compose your findings in a "PowerPoint-like" presentation. You shall present these results in the technical interview. The time allocated to the presentation is of 20 min.*

### **Scenario:**

EnergyInc, is a major utility company providing gas and electricity to corporate, SME and residential customers. In recent years, post-liberalization of the energy market in Europe, EnergyInc has had a growing problem with increasing customer defections above industry average. EnergyInc has asked you to work alongside them to identify the drivers of this problem and to devise and implement a strategy to counter it. The churn issue is most acute in the SME division and thus they want it to be the first priority.

The head of the SME division has asked whether it is possible to predict the customers which are most likely to churn so that they can trial a range of pre-emptive actions. He has a hypothesis that clients are switching to cheaper providers so the first action to be trialed will be to offer customers with high propensity of churning a 20% discount.

### **Your task:**

We have scheduled a meeting in one week's time with the head of the SME division in which you will present our findings of the churn issue and your recommendations on how to address it.

You are in charge of building the model and of suggesting which commercial actions should be taken as a result of the model's outcome. The first stage is to establish the viability of such a model. For training your model you are provided with a dataset which includes features of SME customers in January 2016 as well as the information about whether or not they have churned by March 2016. In addition to that you have received the prices from 2015 for these customers. Of particular interest for the client is how you frame the problem for training.

Given that this is the first time the client is resorting to predictive modelling, it is beneficial to leverage descriptive statistics and visualisation for extracting interesting insights from the provided data before diving into the model. Also, while it is not mandatory, you are

encouraged to test multiple algorithms. If you do so, it will be helpful to describe the tested algorithms in a simple manner.

Using the trained model you shall “score” customers in the verification data set (provided in the eponymous file) and put them in descending order of the propensity to churn. You should also classify these customers into two classes: those which you predict to churn are to be labelled "1" and the remaining customers should be labelled "0" in the result template. You will submit this file with your presentation and your predictions will be scored with area under the ROC curve and Brier score which you shall be discussed during your presentation session.

Finally, the EnergyInc would like to have a view on whether the 20% discount offer to customers predicted to be churned is a good measure. Given that it is a steep discount bringing their price lower than all competitors we can assume for now that everyone who is offered will accept it. According to regulations they cannot raise the price of someone within a year if they accept the discount. Therefore, offering it excessively is going to hit revenues hard.

**Table 1** describes all the data fields which are found in the data. You will notice that the contents of some fields are meaningless text strings. This is due to "hashing" of text fields for data privacy. While their commercial interpretation is lost as a result of the hashing, they may still have predictive power.

Table 1:

| Field Name                          | Description   |
|-------------------------------------|---|
| contract_id                         | contract id   |
| company_category                    | category of the company's activity                                |
| campaign_code_elect                 | code of the electricity campaign the customer last subscribed to  |
| channel                             | code of the sales channel   |
| elect_cons_at_12m                   | electricity consumption of the past 12 months                     |
| gas_cons_at_12m                     | gas consumption of the past 12 months                             |
| elect_cons_last_month               | electricity consumption of the last month                         |
| contract_activation_date            | date of activation of the contract                                |
| contract_end_date                   | registered date of the end of the contract                        |
| contract_fist_act_date              | date of first contract of the client                              |
| contract_last_mod_date              | date of last modification of the product                          |
| contract_renewal_date               | date of the next contract renewal                                 |
| fcst_bill_baseline_elect_next_month | forecasted electricity bill baseline for next month               |
| fcst_bill_baseline_elect_cal_year   | forecasted electricity bill baseline for calendar year            |
| fcst_bill_baseline_elect_12m        | forecasted electricity bill baseline for 12 months                |
| fcst_consumption_elect_next_month   | forecasted electricity consumption for next month                 |
| fcst_consumption_elect_12m          | forecasted electricity consumption for next 12 months             |
| fcst_consumption_elect_cal_year     | forecasted electricity consumption for next calendar year         |
| fcst_current_discount               | forecasted value of current discount                              |
| fcst_bill_meter_rent_12m            | forecasted bill of meter rental for the next 12 months            |
| fcst_price_period_1                 | forecasted energy price for 1st period                            |
| fcst_price_period_2                 | forecasted energy price for 2nd period                            |
| fcst_price_power_period_1           | forecasted power price for 1st period                             |
| client_got_gas                      | indicated if client is also a gas client                          |
| current_paid_consumption            | current paid consumption  |
| gross_margin_power_elect            | gross margin on power subscription                                |
| net_margin_power_elect              | net margin on power subscription                                  |
| number_active_products              | number of active products and services                            |
| total_net_margin                    | total net margin  |
| customer_antiquity_in_years         | antiquity of the client (in number of years)                      |
| code_elect_campaign_first_subs      | code of the electricity campaign the customer first subscribed to |
| subscribed_power                    | subscribed power  |
| reference_date                      | reference date  |
| price_energy_1st_period             | price of energy for the 1st period                                |
| price_energy_2nd_period             | price of energy for the 2nd period                                |
| price_energy_3rd_period             | price of energy for the 3rd period                                |
| price_power_1st_period              | price of power for the 1st period                                 |

|                          |   |
|--------------------------|---|
| price_power_2nd_period   | price of power for the 2nd period             |
| price_power_3rd_period   | price of power for the 3rd period             |
| has_churned_over_next_3m | has the client churned over the next 3 months |