**Case Interview**

As a next step, we would like you to prepare 5 to 10 slides around the below case and using the attached information, and present us your presentation during the call interview

A leading telecoms operator in Indonesia would like to measure effectiveness of their campaigns. In particular, they want to answer;

* Are my campaigns effective overall?
* What type of customers should I target?
* Are there customers that I should not target?
* How can I maximize response rates of my campaigns?
* How can I optimize my campaigns?

They would like you to analyze the provided campaign data and come up with an analytical solution to their problems.

Data Description

* Main Data Sheet
  + This is the data you should work on, analyze the data, build your models and report your results
  + There is one campaign, having 5k customers
  + Customers are divided into treatment and control groups
    - If customer is in treatment (Treatment=1) => Offer the product
    - If customer is in control (Treatment=0) => No offer
  + Response shows whether customer purchased the product or not
    - Response=1 : Purchased
    - Response=0 : Not Purchased
  + It is clear that customer may purchase the product even if he is not targeted, or he may not purchase the product even if he is targeted
* Test Data Sheet
  + Based on your model developed in main data, you should reflect your predictions here. As you can see, response field is blank (and marked as yellow) and you need to fill the blanks with your model’s predictions
  + Once you submit your response predictions on “Test Data”, we will evaluate your predictions with realized numbers and find your model’s actual performance

Remarks:

* Test Data Sheet is only for us to evaluate your results objectively
* All modeling work should be done on the main data. In other words, you still need to divide your main data into training, validation or out-of-sample data sets. Once you finished your modeling and come up with a final model, as an extra step, we want you to predict the responses in Test Data Sheet

Hints:

* We don’t impose any constraint on how you model the problem or what types of ML algorithms you use. However, we strongly suggest you to consider uplift modeling (or incremental modeling)
  + https://en.wikipedia.org/wiki/Uplift\_modelling
* Evaluate your model in terms of Bias-Variance tradeoff
* Evaluate your model in terms if over/under-fitting and model complexity
* Consider sample sizes (sufficient for statistical inference?)
* Discuss pros and cons of your solution approach
* List your assumptions clearly if any
* Mention what kind of other information you need to optimize your campaigns
* Based on provided data, are you able to answer all questions asked in the beginning?
* Assume that audience consists of both data scientists and business owners

Should you have any questions with regards to this, please feel free to contact me directly.

Contact Detail: [mohamedm@xl.co.id](mailto:mohamedm@xl.co.id)