**Candidate technical review questionnaire**

1. **Business ask**

You are in a meeting with a product owner, they just launched an additional customer care center of excellence (CoE), and have opportunity to cross-sell/upsell customers with additional products. They need your help in identifying which customers to pitch to and what product.

At your disposal is historical data for the last 12 months of customer upgrades, their demographic and transaction data. The business has 4 products - video, data, home security and Voice. A customer might be subscribed to a single product, a combination of two or three products or have all 4 products.

a). Describe how you would go about providing a solution to the business? What modeling approaches would you use? How would you measure the effectiveness of your solutions?

* Find correlations between customer info (demographics, transactions, original product subscription) and the upgrade
* See if there is a “signal” that the customer may upgrade soon using the historical data. Plot transactions over time and see if there is any specific pattern in transactions that preceed an upgrade. If we can detect a pattern, this might help us define some input variables for a model to predict when a customer is likely to upgrade.

Classification to predict likelihood of upgrade to ‘x’ product

Train the model with a large portion of the historical customer data (for example, 70% of customers). Score the remaining customers (i.e. remaining 30%) and measure the performance with metrics such as accuracy score (add definition), AUC, Lift chart.

b). Now assume, the business partner has prepared a pipe delimited dataset for you. The dataset contains 128 variables and 300,000 records. The dataset contains records of customer that were serviced by the CoE in the last 9 months, and the analyst created a binary indicator (1 = YES, 0 = NO) of which customers were cross-sold as the target variable. The remaining independent attributes are based on the snapshot view of the customer’s data as of the time of service at the CoE. What modeling approaches would you use? What would your output to the business look like? How would you measure the effectiveness of your solutions? Based on your analysis, what recommendations and or insights do you have for the business?

1. **Logical / process questions**

If you randomly type a 6 digit number on a note, what is the probability that you can see the same number if you flip your note upside down? How would you explain your answer to a 6 year old?