Assignment Statement – Change of Primary Care Physician

An insurance provider (US based) offers health insurance to customers. The provider assigns a PCP(primary care physician) to each customer. The PCP addresses most health concerns of the customers assigned to them. For various reasons, customers want change of PCP. It involves significant effort for the provider whenever the customer makes a change of PCP.

You will find a subset of the insurance provider data along with PCP changes. The provider likes to understand why are members likely to leave the recommended provider. Further, they like to recommend a provider to them that they are less likely to leave.

The dataset consists of following fields:

Id: Column identification field

OUTCOME: Member changed to his/her preferred primary care provider instead of auto assigned to.

0: Member keeps the auto assigned provider.

1: Member changed to this provider by calling customer service.

Distance: Distance between member and provider in miles.

Visit\_count: Number of claims between member and provider.

Claims\_days\_away: Days between member changed to / assigned to the provider and latest claim between member and provider.

Tier: Provider Tier from service, value 1, 2, 3, 4. Tier 1 is highest benefit level and most cost-effective level.

Fqhc: Value 0 or 1

1 : Provider is a certified Federally Qualified Health Center.

Pcp\_lookback: Value 0 or 1.

1: the provider was the member primary care provider before.

Family\_Assignment: Value 0 or 1.

1: The provider is the pcp of the member in the same family.

Kid: 0 or 1.

1: Member is a kid. (under 18 for state of New York)

Is\_Ped: Value 0 or 1.

1: Provider is a pediatrician.

Same\_gender: Value 0 or 1.

1: provider and member are the same gender.

Same\_language: Value 0 or 1.

1: Provider and member speak the same language.

Same\_address: Value 0 or 1.

1: The re-assigned provider has the same address as the provider pre-assigned.

Aim of the assignment is to

* Building a Predictive Model (Which features decide attrition?)
* Evaluate the model.
* Refine the model, as appropriate

The students need to

1. Select a method for performing the analytic task
2. Preprocess the data to enhance quality
3. Carry out descriptive summarization of data and make observations
4. Identify relevant, irrelevant attributes for building model.
5. Perform appropriate data transformations with justifications
6. Generate new features if needed
7. Carry out the chosen analytic task. Show results including intermediate results, as needed
8. Evaluate the solutions
9. Look for refinement opportunities

Following are some points for you to take note of, while doing the assignment:

* Prepare a Report that will supplement the submitted codebase.
* State all your assumptions clearly
* List all intermediate steps and learnings
* Make the report structured and readable.