Greener State Readme

Objective

* Increase customer involvement in the “Greener State” movement.
* Predict the propensity that a new customer would join “Greener State” by analyzing existing customers already in the movement.

Data Source

The data used in the project is from the following sources:

1. Gs: Greener State Account list.xlsx
   * Number of enrolled customers: 1,722
   * The following Gs variables were used



1. Acxiom: acxiom201811\_processed.csv

* Processed from raw Acxiom 2018 Nov data
* The following Acxiom variables were fed into the model.

'ky\_ba','p\_Env\_Issues', 'p\_technology', 'p\_green\_living', 'Comfort\_Consumption',

'Green\_Affinity', 'Info\_Action\_Cap', 'acxiom\_seg', 'Tech\_Prop\_Scr', 'Segment', 'division'

Data Modeling

1. Merged Acxiom and Gs data sources by ky\_ba (Enrolled label: = 1 if ky\_ba in Gs , = 0 if ky\_ba not in Gs ).
2. Performed feature engineering: selected correlated variables, changed categorical data to dummy variables, permuted missing data by filling na with 0.
3. Treated unbalanced data: Oversampling from minority labeled category
4. Performed Random Forest modelling: split data into training set and testing set; fit random forest model with training data; applied fitted model to testing data

Results

Results of the analysis are provided in the following files:

* greener\_state\_propensity.csv

variables: ky\_ba, greener\_state\_propensity

Next Steps

1. Update the model with New Enrolled Account List and Acxiom data from Nov 2018 onwards.
2. Increase the number of features on the customer for better prediction. Additional feature could include first experience with the movement, enrolment date, contribute amount, energy consumption, and participation in other movement.
3. Refine model by performing hyper parameters tuning and cross validation.