AI-LYZE: Technical Documentation

Causal Average Treatment Effect

Various AI algorithms are used to estimate the average treatment effect of the data. Broadly, for each treated individual, the algorithms seek to identify a custom group of untreated individuals that share similar characteristics (e.g., age, gender) to compare him/her to. The overall impact of the program is then calculated by comparing the outcomes of these two groups. Examples of such algorithms include:

1. Linear regression
2. Tree-based methods
    a. Causal Forest
    b. Random Forest
    c. XGBoost
3. Propensity score methods
    a. Stratification
    b. Matching
    c. Weighting
4. Double machine learning methods
    a. Double machine learning
    b. Linear doubly robust learning
5. Meta-learners
    a. T-learner
    b. X-learner

For example, under 4a double machine learning, the algorithm (a) predicts the outcome using the independent variables and (b) predicts the treatment using the independent variables. To estimate the average treatment effect, the algorithm then predicts the residuals from (a) using the residuals from (b).

To determine which algorithm should be used, various robustness checks are performed. Examples of such checks include:

Robustness check	Ideal Outcome
Replace the true treatment with an independent random variable	This should cause the causal estimate to go to zero
Replace the true outcome with an independent random variable	
Add an additional, unobserved common cause variable	This should not affect the causal estimate too much
Add an independent random variable as a common cause	
Perform the analysis on a random subset of the data	
Perform the analysis on bootstrapped samples of the data	

The algorithm that is most robust to these different robustness checks is selected and the average treatment effect for that algorithm is shown, using bootstrapped standard errors.

If none of the algorithms are robust to these checks, then it is likely that the underlying graphical model is wrong, and the user will then be prompted to re-specify the model or collect more data.

Heterogenous Treatment Effect Algorithms

Various algorithms are used to estimate the heterogeneous treatment effects of the data. Examples of such algorithms include:

1. Double machine learning methods
   a. Linear double machine learning
   b. Sparse linear double machine learning
   c. Doubly robust learning
   d. LASSO double machine learning
2. Meta-learners
   a. X-learner
   b. S-learner
   c. T-learner
   d. R-learner
   e. Domain adaptation learner

For each algorithm, the residuals are calculated and an "R-squared" is calculated. The algorithm with the highest R-squared is selected. The heterogeneous casual impact for that algorithm is shown, using bootstrapped standard errors.

To identify which sub-groups were impacted the most/ least, the Shapley values for the selected algorithm are calculated and the variables with the highest variable importance are selected.

Selected Technical Outputs of the Model

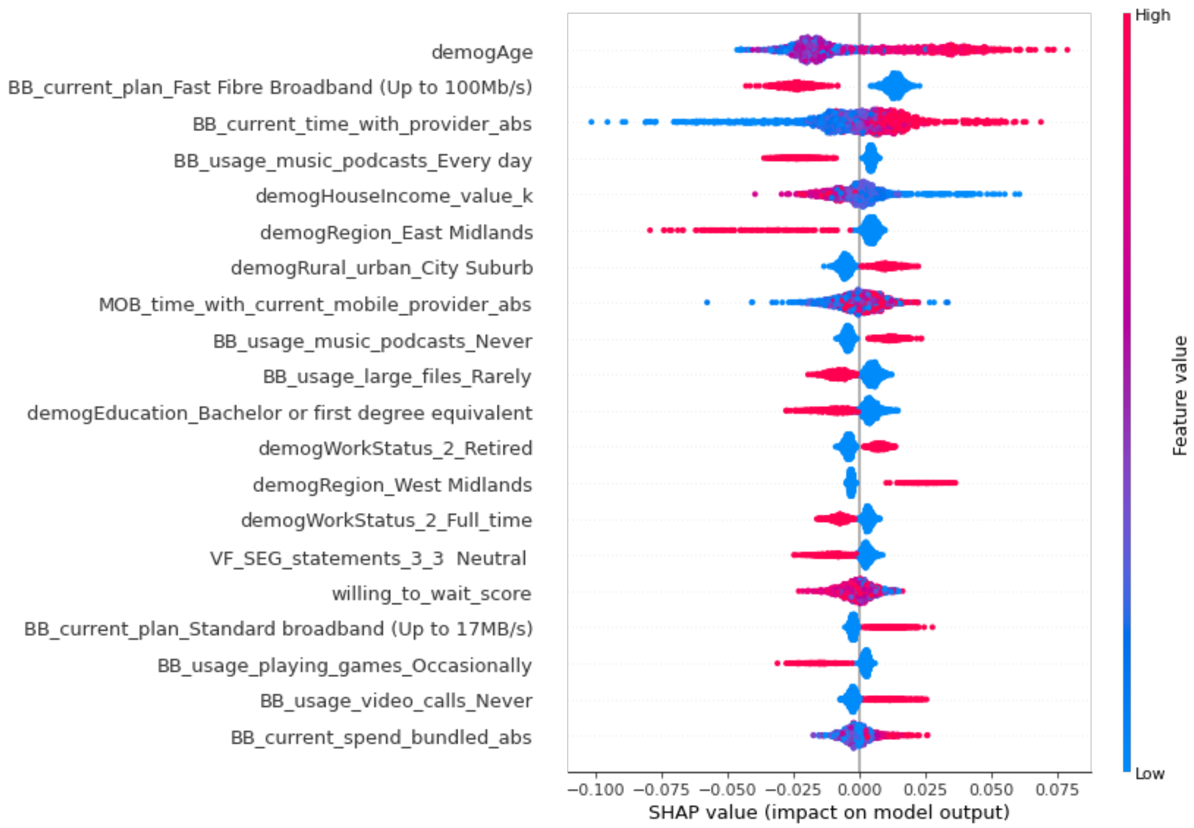
Average Treatment Effect

The average treatment effect model developed for “VF_VoIP_data_for_MIT.csv” passed our robustness checks. Example of such checks include:

Robustness check	Passed?
Replace the true treatment with an independent random variable	Yes
Replace the true outcome with an independent random variable	Yes
Add an additional, unobserved common cause variable	Yes

Heterogeneous Treatment Effect

The variable importance of the heterogenous treatment effect model developed for “VF_VoIP_data_for_MIT.csv” are as follows:



The variables with highest variable importance, namely “demogAge”, “BB_current_time_with_provider_abs”, and “BB_current_plan_Fast Fibre Broadband (Up to 100Mb/s)” were selected.