Exponential Knowledge:
Individual-Level Model

0ptimus

May 14, 2020

Context

From March 27th until April 3rd, our research group fielded two survey experiments aimed at gauging respondents' knowledge of exponential growth patterns. Our work using these two experiments is presented in "The Coronavirus Exponential: A Preliminary Investigation into the Public's Understanding."[1] This document discusses a further line of analysis conducted using these data: a general population look-alike model.

The purpose of the model is simple: to identify pockets (whether geographic or demographic) of individuals who could most benefit from additional public service announcements regarding the potential spread of disease. In this phase of the pandemic, this becomes especially important in communities where no large outbreak has yet been reported.

Model Details

For the purposes of creating a nationwide feature set from which to predict exponential knowledge and for matching respondents to such a nationwide file, we use the L2 commercial file (current as of March 28, 2020). Respondents are matched to the file using a series of exact match filters and, where appropriate, fuzzy matching algorithms.

After matching, the responses of respondents are fed into five different modeling pipelines, including two iterations of XGBoost with slightly different parameters, two iterations of Random Forest, and one logistic regression model. The dependent variable is a binary true/false of whether the respondent was able to correctly identify an exponential trend in our "Understanding Exponential Growth" experiment (Section 3.2). The best model, chosen with ROC AUC as the main evaluation criterion, is then used as the basis for scoring the full nationwide file.

The final outputs of the model are individual-level predictions on the likelihood that an individual is not knowledgeable about exponential trends. From there, aggregated subgroup analysis as well as cutting universes of individuals becomes possible.

Contact

Work on this model is ongoing. For more inquiries about this project, please reach out to Scott Tranter at scott@0ptimus.com.

[1] Podkul, A., Vittert, L., Tranter, S., & Alduncin, A. (2020). The Coronavirus Exponential: A Preliminary Investigation into the Public's Understanding. Harvard Data Science Review. Retrieved from https://hdsr.mitpress.mit.edu/pub/imsfxwvi