CROWDSOURCING PREDICTORS FOR MODELING BEHAVIORAL OUTCOMES

ABSTRACT -- The advancement in the field of digitized information has led to huge improvements in computational tools, data mining methods and in analysis and classification of data. This has taken over the classical approach to science where based on a hypothesis, an experiment is designed, and then data is collected and analyzed. In this work, an alternative to the classical approach is presented where the task of finding correlates for modeling a human behavioral outcome is accomplished through crowdsourcing. This work implements the aforementioned idea by building a website where the crowd is responsible for both creating a questionnaire by posing new questions and providing responses for questions created by the individual and predecessor users of the site. It was also tested that whether the crowd can be motivated to participate in the experiment without financial incentives but with a sense of competition. It was found that the crowd was able to successfully generate many of the known predictors of Body Mass Index.

The results show that:
  i. Non-domain experts were able to generate useful predictors of a human behavioral outcome.
  ii. Users can be motivated to contribute to online experiments without any financial incentive.
  iii. Competition may be useful in improving performance on such a task.

It is observed that many of the predictors generated were non-linearly related to the outcome variable. This work yields more accurate results when a nonlinear model is implemented in comparison to a linear model. The method can be implemented to crowdsource known (and possibly) unknown predictors of many other behavioral outcomes.