Child Abuse & Neglect 46 (2015) 182–189

Contents lists available at ScienceDirect

Child Abuse & Neglect

Research article

Predictive analytics and child protection: Constraints and opportunities

Jesse Russell

National Council on Crime and Delinquency, 426 South Yellowstone Drive, Suite 250, Madison, WI 53719, USA

ARTICLE INFO

Article history:
Received 4 October 2014
Received in revised form 8 May 2015
Accepted 27 May 2015
Available online 2 July 2015

Keywords:
Predictive analytics
Risk assessment
Child protection
Decision making

ABSTRACT

This paper considers how predictive analytics might inform, assist, and improve decision making in child protection. Predictive analytics represents recent increases in data quantity and data diversity, along with advances in computing technology. While the use of data and statistical modeling is not new to child protection decision making, its use in child protection is experiencing growth, and efforts to leverage predictive analytics for better decision-making in child protection are increasing. Past experiences, constraints and opportunities are reviewed. For predictive analytics to make the most impact on child protection practice and outcomes, it must embrace established criteria of validity, equity, reliability, and usefulness.

Introduction

Increases in data quantity and data diversity, along with advances in computing technologies, have made predictive analytics a powerful tool for helping to guide decision making. IBM has suggested that across the globe, 2.5 quintillion bytes of data are created every day. The volume of new data creation is so great that ninety percent of data in the world today was created in the last two years (IBM). In response to these increases in data availability, predictive analytics applications have become common in business settings and are increasing in social services settings.

As some look to leverage the potential of data and predictive analytics, it is important to consider the opportunities and the constraints for using predictive analytics in child protection decision making. The use of data to inform child protection decision making is not new; neither is the use of data and statistical modeling to estimate likelihoods of particular events and assign predictive scores to child protection agency clients. The question is how the emergence of predictive analytics might inform, assist, and improve the use of data to guide decision making in child protection.

Decision Making in Child Protection

Child protection agencies ask "who" questions, "how" questions, and "why" questions. Who questions are about whom the agency should serve, allocate resources to, and target interventions to. Who questions are about which children and families the agency should be most concerned about, who might be appropriate for a particular intervention, or who might be the best opportunity for prevention. Who questions help child protection systems identify which system-involved families they should focus on.

These questions also help answer "who not?" questions. Given all the families who are referred to the system, which families are unlikely to experience further maltreatment, or which families would not benefit from a particular intervention?

In contrast, "how" questions are about how child welfare agencies should serve children and families, the programs they should develop, or the practices they engage in. How questions can be specific to a particular child or family, a particular

http://dx.doi.org/10.1016/j.chiabu.2015.05.022

program, or general to agency practice. How questions could include things about how to develop an effective service plan for a family, whether a specific program is effective, or whether an agency should change its practice or policies.

"Why" questions, which are most commonly addressed by researchers, are about the causes and consequences of child protection involvement. These questions ask why some families might be more likely than others to experience maltreatment, what drives outcomes, or what are the causes of child welfare outcomes?

Child protection agencies have a mandate to respond to and prevent future child abuse and neglect. Though agencies receive many reports of child maltreatment, no agency has the resources to investigate or serve every family mentioned in every report. Moreover, the interventions child protection agencies offer are not always welcome or entirely benign. It is central to the mandate of child protection agencies to decide which families and children to serve—a "who" question.

Predictive models directly relate to these types of questions to inform specific decision points in child protection. For example, a question about which families might be more likely to experience future maltreatment (a family's risk of future child abuse or neglect) can inform the decision of whether to open a case for services. Informed by the outcome of a risk assessment that provides an estimate of the likelihood of future maltreatment, the agency may end its involvement with the family or decide to open a case for services. In California, for example, low- and moderate-risk families are generally recommended for closure unless safety threats remain unresolved, while high- and very high-risk cases are recommended for case opening (Wicke Dankert & Johnson, 2014).

The focus of this article is on these types of decisions—responses to who questions—that can be informed by predictive models.

Predictive Analytics

An emerging approach with the potential to inform decision making in child protection, predictive analytics looks at the past experiences of an organization to estimate the likelihood of future events. Predictive analytics looks at that past by using computer algorithms to sort through an organization's data to produce and "train" (or shape) a model that can then estimate likelihoods of particular events and assign predictive scores to the organization's clients.

Predictive analytics can include a broad set of statistical and analytical tools that identify trends, relationships, and patterns within data that can be used to predict a future event or behavior. Predictive analytics as a broad category can include the concepts and methods associated with "big data," data mining, machine learning, classification and regression trees (CARTs), and random forest modeling, among others.

Standards for Judging Predictive Models

The standards by which predictive models in child protection should be judged are well-established. They should be valid, reliable, equitable, and useful (D'Andrade, Austin, & Benton, 2008). Even if the methodology by which they are developed and used in practice varies and changes, these are the standards by which they must be evaluated.

Validity, in general, refers to how well a test or task matches the attribute or the domain of the knowledge we wish to assess. Validity is about whether the test measures what it is meant to measure. While the concept of validity is clear, actual measurement of validity can be challenging. Most commonly, validity is statistically assessed through the receiver operating characteristic (ROC) or the area under the ROC curve (AUC) (see, inter alia, Fogarty, Baker, & Hudson, 2005).

The ROC is a graphical plot of the true positive rate against the false positive rate (Fogarty et al., 2005). In a binary test (with one "positive" outcome and one "negative" outcome) the true positive rate represents how many correct positive results are achieved among all positives in the sample—or how often the test says the result is positive when the outcome was actually positive. The false positive rate, on the other hand, represents how many incorrect positive results are achieved among all negatives in the sample—or how often the test says the result is positive when the outcome was actually negative. This concept allows analysts to weigh the costs and benefits of a particular test along with the trade-offs between having fewer false positives and more true positives.

Calculating accuracy from the ROC is done simply by adding the number of true positives and the number of true negatives as a fraction of the total sample. In this way, ROC accuracy represents how often the test produces a correct result.

A single measure that can be derived from the ROC is the AUC, which represents the probability that for a randomly drawn pair (one from the positive group and one from the negative group), the test will rank or score the positive case higher than the negative case (assuming that positives are higher on the scoring scale). In other words, the AUC measures the percentage of randomly drawn pairs for which the test correctly classifies both cases.

Interpretation of the AUC in one sense is easy: High AUC values represent more accurate classification. The AUC is often seen as the standard measure for comparing validity across classification models because it is a single metric that can reduce the appearance of subjectivity (Hand, 2009; Lobo, Jiménez-Valverde, & Real, 2008). The potential of the AUC is that two models can be directly compared in a sort of validity competition—models with higher AUC scores might be understood as better classifiers or as more valid (see Rice & Harris, 2005). The literature on AUC and potential alternatives is robust (see, for example, Hand, 2009, or Hanczar et al., 2010).

While the greatest endorsement of the AUC may be that it performs better than other single measures of model validity (Bradley, 1997), trouble in interpreting the AUC can stem from reducing the ROC to a *single* measure rather than using the ROC framework to examine multiple trade-off options (Powers, 2012). As far as single measures of model validity go, the