Public Safety Risk Assessment CLEARINGHOUSE

Forecasting Murder Within a Population of Probationers and Parolees: A High Stakes Application of Statistical Learning

The Policy Digest Series introduces emerging and promising scholarly research on risk and needs assessments to a broad audience.

This policy digest summarizes and discusses policy implications of a 2009 study of machine learning applied to the community supervision context:

Berk, R., Sherman, L., Barnes, G., Kurtz, E., & Ahlman, L. (2009). Forecasting murder within a population of probationers and parolees: a high stakes application of statistical learning. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *172*(1), 191-211.

In 2004, the Crime and Justice Institute and the National Institute of Corrections released a set of eight principles for evidence-based practices in community corrections. Among them was implementing the "risk-need-responsivity" model, which states that resources should be targeted toward those with the highest risk of reoffending, interventions should be targeted to criminogenic needs, and interventions should be responsive to individual characteristics such as gender and developmental stage.

Many jurisdictions now assess risk of recidivism as part of their standard practice in community corrections. Yet as Berk and others note in their 2009 study, "Forecasting murder within a population of probationers and parolees: A high stakes application of statistical learning", risk is typically not measured based on the seriousness of the offense, even though more serious offenses often are larger public concerns and have higher public costs. This makes it challenging for public safety officials to tailor crime prevention strategies to target the most pressing public safety challenges.

Berk and his colleagues wanted to demonstrate how risk assessment can help target resources to the highest-need individuals. They developed a tool to assess the risk of being charged with homicide or attempted homicide for individuals on community supervision in Philadelphia. The Philadelphia Adult Probation and Parole Department (APPD) agreed to create a unit to provide more intensive and specialized supervision and services for the people the tool predicted to be at high risk of committing or attempting to commit murder over a two-year period.

To develop the tool, Berk and his colleagues used a technique called "machine learning", a type of artificial intelligence that allows computers to make predictions or decisions by picking out complex patterns in data without being explicitly programmed. Specifically, Berk used an algorithm called "random forests," one of many algorithms that fall under the machine learning umbrella. Other demonstrations of machine learning in the criminal justice field have often found that it is more accurate than the statistical methods traditionally used to predict risk.

Machine learning is not typically used to identify the direct cause of an event, but to describe the factors that are associated with that event happening. In other words, the goals of Berk and his colleagues' analysis were to examine data about a set of people on community supervision to identify factors associated with people who commit homicide, use those factors to predict the risk levels of a new group of people on supervision, and measure how well the analysis could correctly predict risk of being charged with homicide.

Using information that would be available to APPD at intake, such as age, gender, race, criminal history, and neighborhood characteristics, the tool correctly classified 93 percent of cases. On first glance, the tool seemed to be performing well. However, interpreting the results of predicting rare events such as homicide requires caution—because only a little over one percent of people in the data were charged with murder or attempted murder, even if the model guessed that no one would be charged with murder, it would be right almost 99 percent of the time. Yet that model would not pick up on any of the people who would be charged with homicide, calling into question its utility for preventing serious crimes.

This type of application demonstrates why, even if a risk assessment tool has high overall accuracy, it can also be important to measure how well the model correctly predicts positive cases (i.e., people who are charged with homicide) compared to negative cases (i.e., people who are not charged homicide). Policymakers should consider whether it is more important to correctly predict one category over another depending on which outcome is most important for making a particular policy decision. One way to do this is to consider the costs of incorrect predictions—is the cost of a false negative the same as the cost of a false positive? In this context, APPD determined that the cost of someone being killed due to a false negative result is greater than someone being placed in a unit with more intensive supervision and services than they needed due to a false positive. Because of this, APPD was willing to accept more false positives than false negatives.

After incorporating the decisions about false positives and false negatives, the final analysis correctly classified 93 percent of the people who did not commit a homicide and 43 percent of the people who did. For comparison, an analysis using logistic regression, a more traditional statistical method for developing risk assessments, correctly predicted 99 percent of people who did not commit a homicide but less than one percent of those who did. The machine learning method was able to better predict which people were at high risk of committing homicide.

However, there is a downside to using machine learning algorithms. It can be difficult to identify exactly how much each factor contributes to predicting the outcome, and some algorithms are more transparent than others. The random forests algorithm provides information on variable importance, which estimates how much the overall prediction accuracy would decrease if a variable were not included in the model. For the prediction of homicide charges in the Philadelphia community corrections population, the three most important factors were age at the start of supervision, age at first contact with the adult court system, and number of prior firearm-related convictions. These factors make sense as predictors, lending additional credibility to the tool.

Implications for Practice

So what can APPD's demonstration teach us about how other jurisdictions can incorporate risk assessments and machine learning into community corrections? Although the study provides evidence that machine learning has the potential to improve risk assessment practices within community corrections, good predictions need to be followed by effective interventions to improve public safety outcomes. As Berk and his colleagues note, an evaluation of APPD's high risk unit would be necessary before we could make any conclusions about whether it reduces homicides committed by individuals on community supervision.

Improved risk predictions from machine learning techniques could also be used to inform other community corrections policies and practices beyond the creation of a specialized high risk unit, such as setting supervision conditions or determining appropriate service referrals. Jurisdictions interested in improving their risk assessment practices and using resources more effectively may wish to consider whether machine learning techniques would fit their needs.

References

Bogue, B., Campbell, N., Carey, M., Clawson, E., Faust, D., Florio, K., Joplin, L., Keiser, G., Wasson, B., Woodward, W. (2004). Implementing Evidence-Based Practice in Community Corrections: The Principles of Effective Intervention. Washington, DC: National Institute of Corrections. Available from http://static.nicic.gov/Library/019342.pdf.

The **Public Safety Risk Assessment Clearinghouse** is a one-stop resource that provides practitioners and policymakers with up-to-date and objective information about risk assessment as well as training and technical assistance on its use.

For more information, please visit: https://psrac.bja.ojp.gov/

Suggested citation for this publication:

Tiry, Emily (2017). *Forecasting Murder within a Population of Probationers and Parolees: A High Stakes Application of Statistical Learning (Policy Digest Number 2017-01)*. Washington, DC: The Public Safety Risk Assessment Clearinghouse.

This project was supported by Grant No. 2015-ZB-BX-K004 awarded by the Bureau of Justice Assistance. The Bureau of Justice Assistance is a component of the Department of Justice's Office of Justice Programs, which also includes the Bureau of Justice Statistics, the National Institute of Justice, the Office of Juvenile Justice and Delinquency Prevention, the Office for Victims of Crime, and the SMART Office. Points of view or opinions in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.