Resultant

Reducing Criminal Recidivism Through Advanced Predictive Analytics

Recidivism is measured by criminal acts resulting in rearrest, reconviction, and/or return to prison with or without a new sentence during a three-year period following the inmate's release.

JUNE 17, 2014 NATIONAL INSTITUTE OF JUSTICE

Note: This does not account for or consider offenders that are sentenced to serve in city or county jails. Typically, the State houses inmates sentenced to durations longer than one year.

Indiana taxpayers and the federal government spend over $700 million per year to run the Indiana Department of Corrections. Housing, meals, and medical services equate to roughly $53 per adult offender per day. Of approximately 27,000 offenders in the system, roughly 17,000 are released every year.

This rapid turnover heightens the need for effective rehabilitative services for incarcerated offenders. Reducing criminal recidivism by successfully transitioning offenders into productive citizens not only provides a positive social outcome, the main goal, but can also result in favorable fiscal impacts.

Advanced data analysis techniques make it possible to target the right programs, at the right time, to the right offenders.

THE SITUATION

Studies investigating causes and solutions to recidivism have established that although rates in the US vary among states, all states suffer from similar social and economic consequences. While these studies can be helpful in reforming policy by providing insight into the general effectiveness of programs and services, they rely on assumptions. As a result, the studies lack statistical rigor, which inhibits the ability to make a substantive impact on offenders.

Reducing recidivism requires an in-depth look into when specific subgroups of offenders have the greatest likelihood to recidivate, which suite of programs will optimally reduce their risk of returning to prison, and how to equip policymakers with the information they need to make informed decisions and investments.

The State of Indiana (the State) asked Resultant to work alongside it in the effort to reduce recidivism in Indiana by:

- Understanding for which subgroup of offenders the problem was most pervasive
- Evaluating the effectiveness of programming
- Providing actionable guidance on specific steps to reduce recidivism

The Resultant team established a plan to apply advanced analytical techniques to cross-agency data, which would in turn provide actionable insights for the State.

Subgroups:

A subset of the population being analyzed with specific characteristics such as an offender who is over the age of 50, is in on drug charges, and had no priors.

THE RESULTANT APPROACH: PARTNERSHIP

The Resultant data analytics team leveraged cross-agency data from the State to help tackle the issue of recidivism in a new and innovative way.

DATA DISCOVERY AND ANALYSIS

Resultant aggregated data from disparate systems including the courts, criminal justice institute, and the offender management system. The team worked alongside the State's subject matter experts to analyze the data and understand the insights by applying a generalizable, proprietary algorithm suite, deemed the "Criminal Acts Risk Quantification" tool to highlight relevant information and eliminate less actionable factors. With a full understanding of the data, the team was able to evaluate program effectiveness in reducing recidivism.

PROGRAM PARTICIPATION OPTIMIZATION

The State was interested in identifying specific programs that were effective in reducing recidivism. Of the six programs evaluated, the team was able to identify not only which programs were effective, but for which offender a program would be most effective, given the individual's unique characteristics, background, and criminal history.

Upon determining the optimal program for each offender, the team analyzed the marginal impact of completing a secondary program.

Once the tool is in place, caseworkers will be empowered to identify optimal programming for specific offenders, all based on their characteristics such as age or offense. The tool will better inform policy directors with information on what combination of programs are most effective, what gaps exist in current programs, and the cost effectiveness and redundancy of programs. In addition, the State will be able to project prison populations.

Effective vs. Ineffective Questions

When attempting to reduce recidivism, it is important to start by asking the right questions. The team refrained from focusing on questions that were simplistic, drew their own conclusions, or allowed external factors to influence answers. Instead, questions were tailored to specific offenders.

EFFECTIVE QUESTIONS

- How can we best rehabilitate this offender?
- If this program is applied at this time to this offender, how much of a decrease in his/her probability of recidivism can we expect?
- How can we best reduce recidivism for the offender population by spending $XX?

INEFFECTIVE QUESTIONS

- Which programs are effective?
- Which facility rehabilitates offenders best?

To determine the effectiveness of offender programs, it is imperative that impact be assessed at the individual offender level; not the treatment level.

THE OUTCOME

The "Criminal Acts Risk Quantification" algorithm tool developed by the Resultant data analytics team on top of the SAP HANA© platform allows data to be effectively seen in a new way by:

- Projecting the future risk of recidivism
- Enabling the creation of individualized and optimized programming for offenders
- Better informing policymakers' decisions on important issues like sentencing reform

With the tool, Resultant and the State are able to specify program participation for specific offenders, which will lead to a data-driven understanding of the most effective programs for each inmate to combat recidivating. With a risk score now tied to every offender, generated via the tool, the State can develop thresholds to determine if the projected impact of the program is greater than the cost incurred.

Faced with a limited amount of funding, the State can more effectively align programs and individual offenders to provide the greatest potential for success. Using Resultant's recidivism tool, the State is using data to reexamine the eligibility requirements of each program for offenders. In addition, the State is able to make data-driven decisions about its allocation of programming and funding.

Methodologies

- Exploratory Data Analysis
- Feature Selection
- Advanced Feature Engineering
- Propensity Score Matching
- Counter-Factual Estimation
- Logistic Regression
- Random Forests
- Clustering
- Mixture Modeling
- High-Dimensional Parameterization
- MCMC

METHODOLOGY DEEP DIVE

Analysis Methodology Brief

Causality conceptualizes program impact as the difference in outcomes under treatment and control, only one of which is observed (e.g. offender enrolls in a program and does not recidivate). Estimating program impact is a missing data problem. Specifically, the counter-factual is unobserved, "what would have happened if the offender did not take the program?"

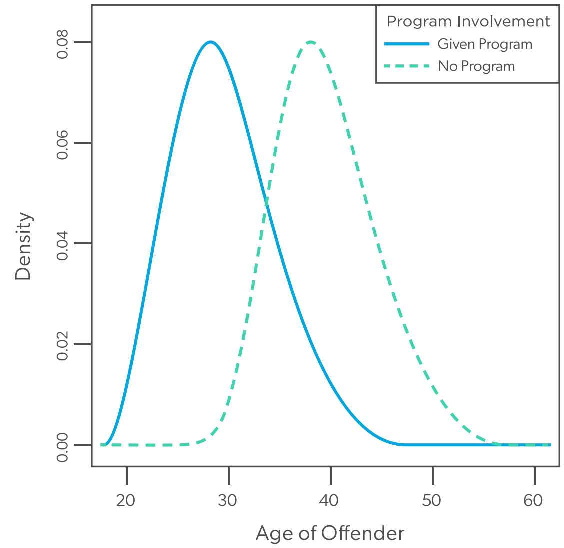
To determine the effectiveness of offender programs, it is imperative that impact be assessed at the individual offender level; not the treatment level. The importance of the individual approach is due to the non-randomized manner of offender participation. When offenders elect to attempt, are court ordered, or advised to take a program, this may lead to differences in characteristics of program participants and non-participants. Failure to control for characteristic differences (prior conviction history, offender classification, age, education, etc.) can lead to biased program estimates due to confounding program impact with characteristic differences.

The graph, shown, portrays two subpopulations defined by the age of the offender. This simple comparison is meant to visualize the challenge and nuances associated with estimating program effects. (In reality, programs often have differences among numerous characteristics.) The population that participated in the program is older relative to the population that did not participate. Given the significant age difference, it is necessary to first identify and control the impact of age, then estimate treatment impact.

If offenders elect to attempt, are court ordered, or advised to take a program, then assignment to the treatment is not random. In observational settings, such as this, balance across treatment and control groups is not guaranteed—resulting in potentially biased estimates of a program's true effect. To recover an accurate estimate of program impact in the presence of non-randomized treatment assignment, Resultant estimates the counter-factual by applying feature selection, propensity score matching, non-linear regression, MCMC, and machine learning techniques.

A VISUALIZING ILLUSTRATION

In this situation, it is difficult to separate program and age effects.



Estimating the Program Effect
E[Yi, received program | received program] –
E[Y, no program | received program]

ABOUT RESULTANT

Our team believes solutions are more valuable, transformative, and meaningful when reached together. Through outcomes built on solutions rooted in data analytics, technology, and management consulting, Resultant serves as a true partner by solving problems with our clients, rather than for them.