Data Driven Organizations, Case Study: COMPAS – Avi Skidelsky

When someone commits a crime the natural reaction is to give them some sort of sentencing through a court trial, typically resulting in serving some amount of time in prison. However, the likelihood of someone being a repeat offender is extremely high, according to a study by the California Department of Corrections and Rehabilitation, more than 65% of those released from California prisons return within three years. These rearrests are reactionary, wouldn’t it be great to be able to predict which criminals would be repeat offenders so that we could put them away for more time to reduce the crime that would happen? That has been the idea of many science-fiction movies and TV shows and is the exact problem that the Correctional Officer Management Profiling for Alternative Sanctions, or COMPAS software tried to solve.

Since the 1970s, criminologists have been trying to predict peoples’ likelihood to commit crimes but had been unsuccessful, especially after they had to stop using racial profiling once it became politically-incorrect to do so. Northpointe Consulting was founded in 1989 by Tim Brennan, a statistics professor, and David Wells, who ran a corrections program in Traverse City, Michigan. Wells had created a prisoner classification system for his jail and wanted to improve upon the Level of Service Inventory, LSI, which had already been developed in Canada as he had found “a fair amount of weaknesses” in it. He wanted a tool that addressed the theories about what causes a crime to be committed which is what led he and Brennan to develop COMPAS at Northpointe.

The purpose of this tool was to assess the likelihood of someone committing a crime which it ranked on a scale of one to ten based on a survey of 137 questions pulled either from a self-answered survey or pulled from prior criminal records. Its intended use was to be a tool for parole recommendation hearings, to provide the judges with an idea of the likelihood of the parolee committing another crime. However, it has started to become frequently used for more intense parts of the judicial process including determining the initial sentencing of a criminal and the price of bail, if at all, for an alleged criminal at pretrial hearings.

As early as 2001, large jurisdictions have been relying on the COMPAS program’s recommendation without fully testing its accuracy, including New York which had started using it for the vast majority of its parole hearings around the state in 2010. Other states including Arizona, Colorado, Delaware, Kentucky, Louisiana, Oklahoma, Virginia, Washington, and Wisconsin have started using it in their sentencings by giving assessments to judges during trial. Wisconsin in particular had rolled it out statewide in 2012 to be used at each step of the prison system from sentencing to parole.

A study in 2012 showed that COMPAS’s assessments were 71% accurate in parole systems, but it failed to evaluate racial differences. Indeed, there was heavy racial bias introduced by COMPAS in that it frequently gave higher risk grades to African Americans and Latinos than to Caucasians, even when Caucasian offenders had prior criminal history and theoretically received higher scores than their minority counterparts. This was pointed out by then US Attorney General Eric Holder who warned that risk scores might be injecting racial bias into the courts when he called for the US Sentencing Commission to study their use. ProPublica conducted a study where they obtained the risk scores for more than 7,000 people in Broward County, Florida in 2013 and 2014 to evaluate how accurate this system actually was. The study proved that COMPAS was extremely unreliable in predicting violent crime, categorizing only 20% of people correctly within the following two years. When it came to predicting crime as a whole, including something as minor as driving without a license, the algorithm was more accurate but still not satisfactory, only 61% correct.

In today’s environment, Northpointe would have an extremely difficult time explaining their software and pitching it to potential customers and even the public at large. In a world where racial discrimination is at the forefront of peoples’ minds, hearing that a risk assessment software is inherently racially biased would be a struggle for Northpointe to overcome. Theoretically, if the software produced extremely accurate results it might be easier to market to the public that the system is not racially biased, and rather it’s just a reflection of society at large. Unfortunately, I don’t believe that their success rate in the 60th percentile would cut it. In order for COMPAS to keep succeeding under heavy scrutiny, it would have to alter its algorithm but as Brennan has pointed out in previous comments, COMPAS does not take race into consideration but rather, there are some questions that are inherently racially biased. These questions are the types that ask about lifestyles and habits that typically belong to distinct minority groups for instance areas of residence, financial standing, family life, or profession. A way to reduce these inherently racial questions is to judge an offender based on criminal record, type of crime committed, and personality traits. This means that the algorithm may have to rely more heavily on the self-answered questions than objective ones, perhaps even introducing a psychological evaluation. However, this may not improve the effectiveness of COMPAS but at least it will bring it to the point that it will have a better public image. Perhaps COMPAS should only be used for parole hearings, which is where it has shown to be the most effective and accurate.