CSC2552: Review 2, Paper 2

Due on January 30

Ashton Anderson

499 words

Thomas Hollis

Paper 2

This fascinating paper by Kleinberg, Lakkaraju et. al. is a brilliant application of machine learning to legal data with the main aim of helping improve judge bail decisions in US courts. The main results of this paper are profound as they show that a gradient-boosted decision tree algorithm can help significantly improve either the offending rate of released suspects or the incarceration rate of suspects for a fixed incarceration rate and crime rate respectively, as well as improve the judges decision making process by revealing their blind spots.

While finding weaknesses in such a revolutionary paper is daunting, one possible significant weakness of this paper is the lack of a large scale custom-made human trial to prove the reported ML algorithm's efficacy rates. Another related weakness is the lack of a clear source to the data and code used in the paper's investigations. However, it seems that both limitations are due to a tradeoff with ethical confidentiality and security (only three small R scripts were released detailing contraction, imputation and model fitting) [1].

In a contrast, a major strength of this paper is the depth and detail in the approach taken. Indeed for every single claim made there is a clear demonstration of how the data was processed and analysed as well as what are the legal and societal ramifications for such results. Extra attention is payed to algorithmic fairness with regards to sensitive features like race, even going as far as attempting to control for latent variable learning by the ML algorithms. Indeed, this task is still a very active area of research and continues to be a major challenge in the field of ML [2]. The main tradeoff here is the large amount of overhead work required but this benefits the research by drastically improving the strength of the paper's conclusions. Some other notable strengths of this paper include clear actionable advice for judges as well as detailed discussions of biases and limitations of the ML algorithm, in particular the limitation of missing data for jailed suspects.

The implications of the paper's results are truly enlightening as they open the door to a new area of algorithmic human-in-the-loop decision making for US courts while also revealing how ML could help reduce the existing burden of the increasingly large volume of cases. This could also drastically improve the lives of hundreds of thousands of people who find themselves (rightly or wrongly) passing through the US judicial system, as it reveals current flaws of judges which can be learned from. The conclusions are stated clearly, with very little compromise required, using strong evidence to back up each claim. Nevertheless, the authors could perhaps have suggested more information on how the results of this paper could be implemented in real life. Indeed, algorithmic tools are already in use in the US criminal justice system such as COMPAS, a highly controversial risk assessment tool used to predict recidivism, which has received widespread criticism from academia [3].

[1] Kleinberg, Jon; Lakkaraju, Himabindu; Leskovec, Jure; Ludwig, Jens; Mullainathan, Sendhil, 2017, "Replication Data: Human Decisions and Machine Predictions", https://doi.org/10.7910/DVN/VWDGHT, Harvard Dataverse, V1

[2] Hardt, M., Price, E., & Srebro, N. (2016). Equality of opportunity in supervised learning. InAdvances in Neural Information Processing Systems(pp. 3315-3323).

[3] Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. Science Advances, 4(1), eaao5580. doi:10.1126/sciadv.aao5580