

Bias in Data-Informed Decision Making

B

Harini Dissanayake

DOT loves data, New Zealand

Paul J. Bracewell

DOT loves data, New Zealand

INTRODUCTION

With vast amounts of data becoming available in a machine-readable format, decision makers in almost every sector are focused on exploiting data and machine learning to drive the phenomena of automated decision making, whilst rapidly dissolving the human oversight in the process. The description of bias in decision making arising from machine learning outlined in this chapter sets to demonstrate the scale of the issue and the value of transparency in decision making that affects the daily life of ordinary humans.

As a practical example, in New Zealand, the extent of algorithm use has already extended to use cases such as a New Zealand Police tool that assesses family violence risk, and a Ministry of Social Development classification tool for identifying school leavers that require the greatest education and employment support. The concern raised by many of the government bodies that adopt operational algorithms is the potential bias in such modelling, with identifying and improving these potential biases a priority in the future (Statistics New Zealand, 2018). This chapter identifies such biases in both government and non-government decision making data, with a special focus in identifying discriminatory bias of operational algorithms and sample selection bias inherent in decision making data.

Behavioral research finds that human decision making is clouded by ‘human errors’. In various contexts, the literature documents that human problem solving tends to systematically differ from the predictions of rational choice models. From the past 50 years of work in behavioral research and decision making, it is well established that human decision making is often not as rational as one might expect. The development of ‘data science’ allows decision makers to detect present day human errors impacting specific decisions in real time. One of the objectives of this chapter is to reveal the extent of human biases found in the decision-making information in present day New Zealand.

Machine learning systems are becoming increasingly prominent in automated decision making in New Zealand. Using systems that are sensitive to the type of bias that results in discrimination, must be undertaken with caution. Given the scale and impact of the bodies that have already adopted machine learning, it is crucial that measures are taken to prevent unfair discrimination through legal as well as technical means. There has been significant effort to avoid and correct discriminatory bias in algorithms while also making them more transparent. In New Zealand, the government recently claimed world first in setting standards for how public agencies should use the algorithms that increasingly drive decision making by officials about every aspect of public life. New Zealand has produced a set of standards; “Algorithm charter for Aotearoa New Zealand”, designed to guide government use of algorithms and to improve data transparency and accountability. It outlines several measures including:

1. how decisions are informed by algorithms:
2. making sure data is fit for purpose by identifying and managing bias

DOI: 10.4018/978-1-7998-9220-5.ch068

3. ensuring that privacy, ethics, and human rights are safeguarded by regularly peer-reviewing algorithms
4. explaining the role of humans in decisions informed by algorithms, and
5. providing a channel for citizens to appeal against decisions informed by algorithms.

New Zealand hopes to set a standard for other nations by leading the way with responsible algorithm oversight and by demonstrating the value of transparency of algorithms that affect daily life. “Many government agencies are already harnessing the power of data to deliver improved public services for New Zealanders – coming up with innovative solutions to complex problems...but as these techniques grow in scale and sophistication, it’s critical that New Zealanders can be confident their data is being handled appropriately, and that proper safeguards are being applied” Minister for Statistics James Shaw said (Shaw, 2019). Dokyun Lee (Business Professor at Tepper School of Business), advises that the art of human interpretation and reasoning cannot be replaced by artificial intelligence - machine learning works best only as a tool (Carnegie Mellon University, 2019). Lee describes “machine learning, if sufficiently transparent, is great for helping humans discern patterns and understand the nuance of complex big data...but inherently, making connections about how the world works and creating hypotheses to find answers that stand the test of time — we (humans) are the only ones who can do that, at least for now.” As deep learning thrived under machine learning, many recognized the potential power of harnessing big data sets, beginning to employ these systems that appeared as a revolutionary panacea. The algorithms on which these systems are based learn by example, rather than rules or data features engineered by humans. The algorithm identifies patterns in the data and arrives at a decision using those patterns. As the machine’s decision-making process occurs outside of human control, the method the machine is learning for drawing its inferences is not readily understood by algorithm users. Consequently, human oversight is critical to evaluate any machine’s decision that could be wrong.

This chapter, which describes bias in decision making, is motivated by the need for approaching any analysis with caution. Fully automated solutions without appropriate governance can be problematic due to the replication of human biases which can be captured and imposed upon the training process of a machine learning models. Bias is omnipresent in machine learning. Decision makers alike must be aware of limitations within the data and proceed with caution. As such, transparency in how data is used to make decisions is vital. Despite the ever-increasing reliance on algorithms in decision making, human oversight is critical (Statistics New Zealand, 2018). This chapter serves to raise awareness of this aspect of machine learning.

DECISION MAKING AS A QUEST FOR RATIONALITY

Behavioral research finds that human decision making is clouded by ‘human errors.’ In various contexts, the literature documents that human problem solving tends to systematically differ from the predictions of rational choice models. For instance, humans tend to focus on irrelevant information (Kahneman, 2013), fall victim to variables that are contextual and situational (Danziger et al., 2011) and go as far as to rationalize the bad decisions made (Harmon-Jones & Mills, 1999).

With vast amounts of data becoming available, decision makers in almost every sector are focused on exploiting data for decision making. However, the volume and variety of data generated exceeds the ability of manual analysis. Data, power of computers and algorithms have accelerated, together powering broader and deeper analyses than previously imagined (Provost & Fawcett, 2013). The convergence of

12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/bias-in-data-informed-decision-making/317521

Related Content

Features Selection Study for Breast Cancer Diagnosis Using Thermographic Images, Genetic Algorithms, and Particle Swarm Optimization

Amanda Lays Rodrigues da Silva, Máira Araújo de Santana, Clarisse Lins de Lima, José Filipe Silva de Andrade, Thifany Ketuli Silva de Souza, Maria Beatriz Jacinto de Almeida, Washington Wagner Azevedo da Silva, Rita de Cássia Fernandes de Lima and Wellington Pinheiro dos Santos (2021). *International Journal of Artificial Intelligence and Machine Learning* (pp. 1-18).

www.irma-international.org/article/features-selection-study-for-breast-cancer-diagnosis-using-thermographic-images-genetic-algorithms-and-particle-swarm-optimization/277431

Survey of Recent Applications of Artificial Intelligence for Detection and Analysis of COVID-19 and Other Infectious Diseases

Richard S. Segall and Vidhya Sankarasubbu (2022). *International Journal of Artificial Intelligence and Machine Learning* (pp. 1-30).

www.irma-international.org/article/survey-of-recent-applications-of-artificial-intelligence-for-detection-and-analysis-of-covid-19-and-other-infectious-diseases/313574

Software for Teaching: Case PROMETHEE

Kenia C. Da Silva F., Coromoto D. M. Raidi Ch., Gilberto J. Hernández G., María J. García G. and José G. Hernández R. (2023). *Encyclopedia of Data Science and Machine Learning* (pp. 1466-1481).

www.irma-international.org/chapter/software-for-teaching/317557

A Survey on Arabic Handwritten Script Recognition Systems

Soumia Djaghbellou, Abderraouf Bouziane, Abdelouahab Attia and Zahid Akhtar (2021). *International Journal of Artificial Intelligence and Machine Learning* (pp. 1-17).

www.irma-international.org/article/a-survey-on-arabic-handwritten-script-recognition-systems/279276

Efficient Closure Operators for FCA-Based Classification

Nida Meddouri and Mondher Maddouri (2020). *International Journal of Artificial Intelligence and Machine Learning* (pp. 79-98).

www.irma-international.org/article/efficient-closure-operators-for-fca-based-classification/257273