*ABSTRACT*

# Modelling and prediction of NZ's population wellbeing using machine learning techniques

Anantha Narayanan T L

**Keywords:** population wellbeing, administrative data, machine learning

The importance of population wellbeing is gaining traction and acknowledgment across the globe. In 2019, the government administered NZ's first ever 'wellbeing budget' after recognising that although NZ had strong economic growth, there were high rates of child poverty, homelessness, and suicide. In NZ, detailed wellbeing measures are available from the General Social Survey (GSS), a biennial survey of approximately 10,000 individuals. Although the GSS sample is considered representative of the NZ population, a major limitation is that the survey sample is not representative of certain groups of the population that may be of high policy interest. For example, the wellbeing of social housing tenants is a key area of government focus, but the number of social housing tenants represented in the GSS survey is sparse. Therefore, it is impractical to explore what/how government policies are associated with wellbeing in these smaller subsets of the population. To overcome this challenge, population-level wellbeing data is required. The government's push for cross-agency data integration over the last several decades led to the development of the Integrated Data Infrastructure (IDI). The IDI is a complex population-level government research database containing anonymised individual response data (microdata) relating to people and households that can be linked longitudinally over time. Extrapolating wellbeing data to the full NZ population may be possible by applying advanced modelling techniques on the IDI data. In this study, I propose to use machine learning algorithms (such as random forests, neural networks) to model and predict various GSS-based wellbeing outcomes for the full NZ population. If successful, the outcome of this study (a population-level measure of wellbeing) would enable researchers to easily incorporate wellbeing measures into IDI-based policy analysis. Ultimately, it would improve our understanding of how the political, social, and economic environment influences the wellbeing and functioning of New Zealanders.