1. Introduction

The goal of this project is to evaluate the sustainability performance of 140 neighbourhoods in the City of Toronto. To this end, the k-means clustering algorithm, a popular (unsupervised) machine learning technique, will be drawn upon using a range of variables pertaining to the three domains of sustainability – social, economic and environmental. The clustering technique will help partition the neighbourhoods into groups (clusters) and identify their distinct characteristic.

The need for such analysis arises from the fact that the evaluation of social, economic and environmental indicators is often performed in isolation from one another. However, analyzing sustainability requires a comprehensive and comparative look at the state of social, economic and environmental well-being of a unit. In this regard, this project will be an attempt to comprehensively examine relevant indicators and propose sustainability profiles by which neighbourhoods can be grouped.

In practical terms, such clustering of neighbourhoods can be useful in:

- setting a benchmark for the sustainability evaluation of neighbourhoods in the future;
- comparing sustainability performance of neighbourhoods in the present;
- identifying suitable neighborhoods for various business projects such as housing development;