

Problem description:

Machine Learning (ML) is a form of AI that enables a system to learn from data rather than explicit programming. Machine learning has become one of the most important topics within organisations that are looking for new ways to leverage data to help the business and gain a new level of understanding. ML algorithms are advantageous since they offer solutions to problems related to the big quantities of different types of data in a speedy way.

Unsupervised learning is best suited when the problem requires a massive amount of data that is unlabelled. Unsupervised learning algorithms segment data into groups of examples (clusters) or groups of features. One of the popular unsupervised algorithms is K-Means. It is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining.

Foursquare is a location technology platform where people comment and rank food sites, coffee sites, malls and parks. I am consulting a customer who is going to choose one of the job options in New York and Toronto. We are trying to analyse the similarities and differences between neighbourhoods in downtown Toronto and Manhattan by using Foursquare location data along with a K-means clustering algorithm. We try to decide by looking most common venues in these areas and try to find a place for his pleasure.