CIND820 D1H - Big Data Analytics Project - F2024

Name: Shankar Ghimire

Student #: 501177872

Program: Data Analytics, Big Data, and Predictive Analytics

Abstract

Childcare facilities play a very important role in managing the job and personal life for many families across the small town to entire province and country. The dataset on ‘Licensed child care facilities in Ontario’ provides comprehensive information on various aspects including their location, language of service, program type etc. The dataset contains 8,184 rows and 18 columns.

The problem being considered for the project is to predict the facility closure, classify the facilities and forecast the demand for childcare services in the various locations. For the prediction, the appropriate algorithms such as logistic regression, neural networks algorithms etc. will be applied on the dataset. Similarly, for the classification of childcare facilities, algorithms such as Decision trees, random forest, K-Nearest neighbors etc. will be applied.

For the classification and prediction tasks on the given dataset, a systematic data analysis approach will be employed. Initially, the entire dataset will go a cleaning process, followed by preliminary analysis utilizing various exploratory data analysis tools. Then, experimental design and model constructions will be undertaken. Finally, the performance of the model will be evaluated with recommendations and conclusions.

References:

Dataset source

<https://data.ontario.ca/dataset/7efd8b4b-cc63-4337-a551-c940a346605b/resource/2b81313c-9ada-4680-abba-4470ec386a2e/download/child_care_facilities_open_data_august2024.xlsx>