CMPT 318 – CYBER SECURITY

DR. UWE GLASSER, FALL 2021

ASSIGNMENT 2.

1. Answers:

Feature scaling is an essential preprocessing technique of input data for machine learning algorithms and beyond. Two widely used feature scaling techniques are normalization and standardization. For each of the following questions explain clearly and concisely:

1. What is feature scaling?

* Feature scaling is a type of modeling that helps raw data that a human can understand be transformed into an input set of data that can be understood by a machine.
* Feature scaling also helps us converge data easier because of proper scaling.

1. Why scaling features of a dataset is necessary?

* Gives us an insight of how the data behaves relative to a number, this not just makes it easier for a person to see certain patterns, normalization or standardization of data makes it easier for a machine to also detect anomalies in the system.
* Clustering algorithms (Machine Learning) will fail as they will be skewed by different (not standardized nor normalized) scales as they rely on distance metrics.

1. What does normalization and standardization do to the data and the noise? Your answers to all three questions should not exceed 2 pages in total but provide technical descriptions including the use of mathematical notation.