**Feature Selection and Reduction**

In machine learning and in data science as a whole, data plays a humongous and pivotal rule. In fact, data is the basic building block of this sphere. And every piece of information, every data can be represented by or so to speak, is represented keeping various features or characterstics in sight. These features are also called dimensions. But all the features that are used to represent a data are not relevant and there can be a huge amount of them.

Some redundant or irrelevant features may also exist in data representation. And therefore all these redundant data need to be eradicated and removed for the proper and efficient application of machine learning algorithms. The higher the number of features, the harder it gets to visualize the training set and then work on it. Sometimes, most of these features are correlated, and hence redundant. This is where dimensionality reduction algorithms come into play. Dimensionality reduction is the process of reducing the number of random variables under consideration, by obtaining a set of principal variables. It can be divided into feature selection and feature extraction.

So technically, Problem of selecting some subset of a learning algorithm’s input variables upon which it should focus attention, while ignoring the rest is the crux of feature selection and reduction. It may also be referred to as Dimensionality Reduction**.**

Main reasons to use feature selection are following: It enables the machine learning algorithm to train faster. It improves the accuracy of a model if the right subset is chosen. It reduces the complexity of a model and makes it easier to interpret

Advantages of Dimensionality Reduction:

It helps in data compression, and hence reduced storage space. It may also reduces computation time and also helps remove redundant features.

But there can also be a disadvantage thatIt may lead to some amount of data loss.

Regards

Ayush Anand

3rd Year Undergraduate

Information Technology

IIEST Shibpur