T/UDOM/2017/15644 SUPERVISED LEARNING REPORT

SUPERVISED LEARNING

Is one of the most commonly used and successful types of machine learning which predict the outcome from a given input.

It takes the form Y=f(x)

Where Y is output and X is input.

TYPES OF SUPERVISED LEARNING PROBLEMS

- Classification
- Regression

REGRESSION

In regression problems the labels are quantitative or continuous in nature. Examples include

- Income in dollars
- Weight in pounds
- Distance in miles

CLASSIFICATION

In classification problems the labels are qualitative or categorical in nature and can be grouped into two or more classes that includes

- Binary labels(yes/no or 0/1)
- Different brands of a product(A,B,C)
- The weather on a given day (rainy, sunny, overcast)

GENERALIZATION, OVERFITTING AND UNDERFITTING

OVERFITTING

This involves building a model that is too complex for the amount of information we have, as our novice data scientist did

UNDERFITTING

This involves choosing too simple a model

GENERALIZATION

Refers to the ability of an algorithm to be effective across a range of inputs and applications.

SUPERVISED MACHINE LEARNING ALGORITHMS

The following are algorithms of supervised machine learning

- kNN
- Decision trees
- Naïve Bayes

- SVM