**Unit-12**

Lecture 2.1.1

|  |
| --- |
| Mapped with CO2,3,4,. |

**Logistic regression**

This type of statistical analysis (also known as *logit model*) is often used for predictive analytics and modeling, and extends to applications in machine learning. In this analytics approach, the dependent variable is finite or categorical: either A or B (binary regression) or a range of finite options A, B, C or D (multinomial regression). It is used in statistical software to understand the relationship between the dependent variable and one or more independent variables by estimating probabilities using a logistic regression equation.

•Logistic Regression was used in the biological sciences in early twentieth century. It was then used in many social science applications.

•Logistic Regression is used when the dependent variable(target) is categorical.

•For example,

•To predict whether an email is spam (1) or (0)

Whether the tumor is malignant (1) or not (0)

•Consider a scenario where we need to classify whether an email is spam or not.

•If we use linear regression for this problem, there is a need for setting up a threshold based on which classification can be done.

•Say if the actual class is malignant, predicted continuous value 0.4 and the threshold value is 0.5, the data point will be classified as not malignant which can lead to serious consequence in real time.

•From this example, it can be inferred that linear regression is not suitable for classification problem. Linear regression is unbounded, and this brings logistic regression into picture. Their value strictly ranges from 0 to 1.

•Output = 0 or 1

•Hypothesis => Z = WX + B

•hΘ(x) = sigmoid (Z)

•

•If ‘Z’ goes to infinity, Y(predicted) will become 1 and if ‘Z’ goes to negative infinity, Y(predicted) will become 0.

Referencebooks

1. Christopher M. Bishop, Pattern Recognition and Machine Learning,Springer, 2006.
2. Ryszard S. Michalski, Jaime G. Carbonell, Tom M. Mitchell, MachineLearning : AnArtificial Intelligence Approach, Tioga Publishing Company, 1983.
3. Mitchell. T, Machine Learning, McGraw Hill, 1997.

**Regression.pptx**