Logistic Regression is a simple machine learning algorithm used to predict the probability of an event occurring. It is mainly used for classification problems, where the outcome is binary (e.g., Yes/No, 0/1, True/False).

**How It Works:**

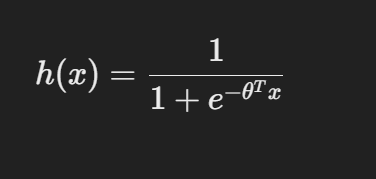
* Logistic Regression finds the relationship between one **dependent** (output) binary variable and one or more **independent** (input) variables.
* It calculates the probability of an event happening using a mathematical function.

**The Role of Parameter C:**

* **C** is a regularization parameter that controls the model’s complexity.
* If **C** is too large, the model becomes too complex and overfits (memorizes) the training data.
* If **C** is too small, the model becomes too simple and underfits (fails to learn enough from the data).
* The best value of **C** is found using a method called **RandomizedSearchCV()**.

**Logistic Regression Formula:**

The model makes predictions using the **Sigmoid function**, which converts any value into a probability between 0 and 1:



Where:

* h(x) is the probability of an event occurring.
* θ represents the model’s learned parameters.

This function ensures that the output is always between 0 and 1, making it useful for binary classification problems.

The **T** in **θᵀ x** represents the **transpose** of the vector **θ** (theta).

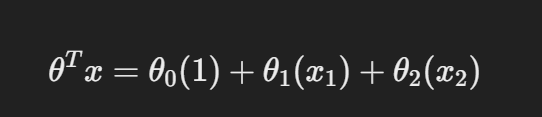
**Explanation:**

* **θ** is a vector of learned parameters (weights) in the model.
* **x** is a vector of input features.
* **θᵀ x** is the **dot product** (matrix multiplication) of the two vectors.

Mathematically, if we have:

* **θ = [θ₀, θ₁, θ₂]** (parameters)
* **x = [1, x₁, x₂]** (features, including a bias term)

Then:



This linear combination is passed into the **sigmoid function** to output a probability.