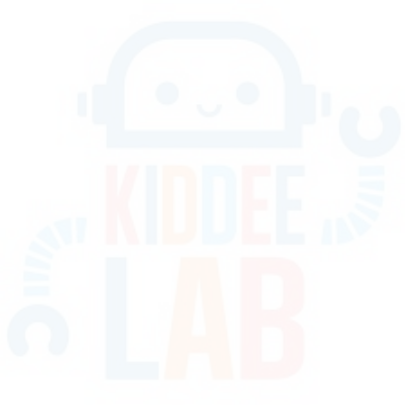
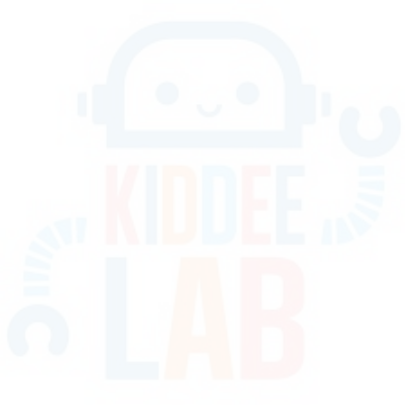
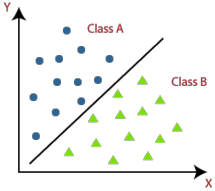
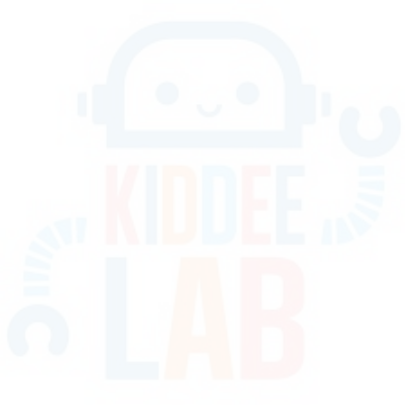
**Supervised Machine Learning Classification**

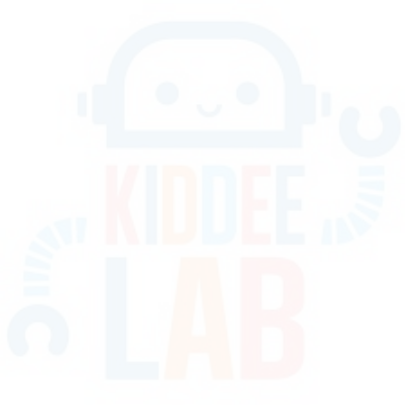
****

What is classification?

- Supervised Machine Learning

- The process of recognition, understanding, and grouping of objects and ideas into preset categories. - It utilizes input training data for the purpose of predicting the likelihood or probability that the data that follows will fall into one of the predetermined categories.

**Logistic Regression**

Logistic Regression 

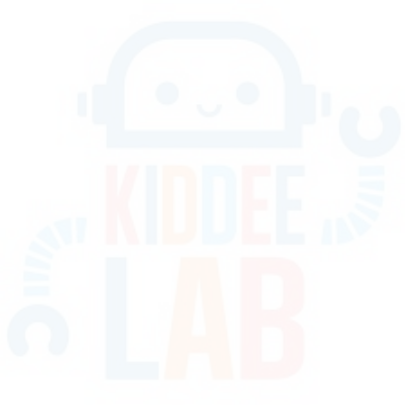
- Statistical Model

- Estimates the probability of an event occurring based on a given dataset of independent variables. - Three types - based on categorical response.

- Binary logistic regression

- Multinomial logistic regression

- Ordinal logistic regression

Binary logistic regression 

● The most commonly used among three logistic regressions ● The most common classifiers for binary

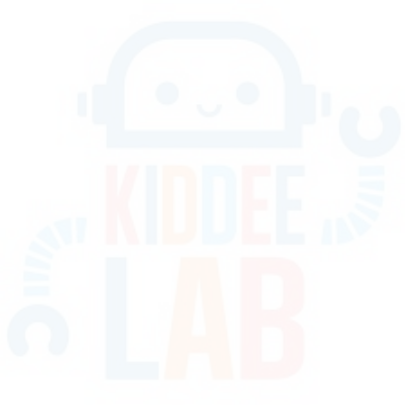
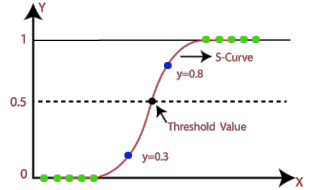
● Only two possible outcomes {0 or 1}.

0: Negative Class 1: Positive Class

● Examples:

- Predict if an e-mail is spam or not spam

- Predict if a person will survive or not in an event

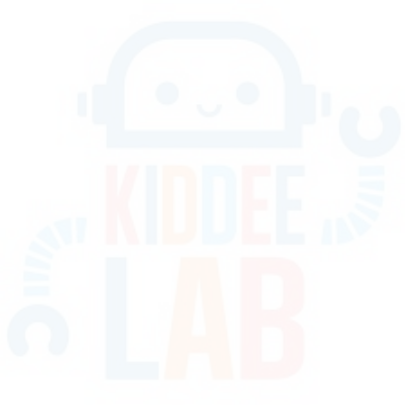
Binary logistic regression Hypothesis Representation: 0 <= hθ(x) <= 1 

"x” is input and “y” is output.



Threshold classifier output, hθ(x) at 0.5 If hθ(x) >= 0.5, predict y = 1

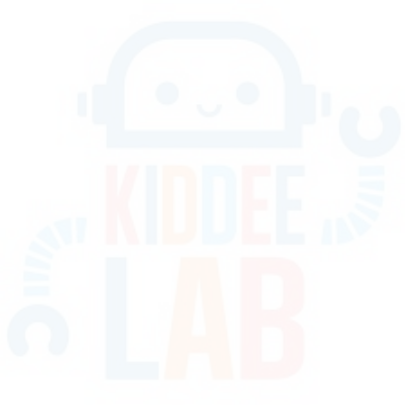
If hθ(x) < 0.5, predict y = 0

Binary logistic regression 

Sigmoid Function 

Logistic Function

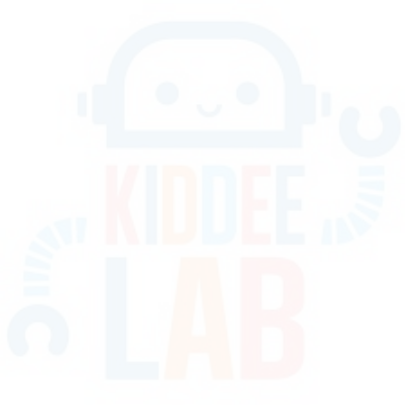
A function that map between 0 and 1

Multinomial logistic regression 

● The dependent variable has three or more possible outcomes; however, these values have no specified order.

● Example:

- Predict what genre of film a moviegoer is likely to see to market films more effectively. A multinomial logistic regression model can help the studio to determine the strength of influence a person's age, gender, and dating status may have on the type of film that they prefer. The studio can then orient an advertising campaign of a specific movie toward a group of people likely to go see it.

Multinomial logistic regression 

Input ~~Class 1~~ No

Class 2

No

Class 3

Yes

Yes Yes

Output

Multinomial logistic regression 

● It uses softmax function.

● Softmax function determines the class with highest probability.

● The class that has highest probability value is the winner and it is the output label of the input data.

Example:

Input is an image of a car, the possible results are as follows.

Bicycle: 0.15

Truck: 0.35

Car: 0.50

Total value of the probability of all label = 1.00

The car has highest probability value. Therefore, it wins.

The output label is the car.

Multinomial logistic regression

It computes the scores. 

k is the class (label of each type)

θ is the model parameter

score

**Soft Max function**

x is the feature values 

σ is the probability that (x) belong to class k given the earlier score

K is the total number of score.



Ordinal logistic regression 

● It is leveraged when the response variable has **three or more possible outcome**, but in this case, these values do have a **defined order**.

● Examples:

- Grading scales from A to F

- Rating scales from 1 to 5

When to apply Logistic Regression 

Almost always, as a baseline though! Logistic Regression make an assumption based on linearity and as long as your data is approximately linear, Logistic Regression work fantastic. There are also some clear advantages:

Advantages:

● They are quite fast for both training and prediction

● They have very few (if any) tunable parameters

● Descent algorithms works well with Logistic Regression

Disadvantages:

● The only problem of Logistic Regression lies on its limitation of linearity.

Classification Performance Measurements Parameters 

● **TN** / **True Negative**: 

○ when a case was negative and

predicted negative

● **TP** / **True Positive**:

○ when a case was positive and

predicted positive

● **FN** / **False Negative**:

○ when a case was positive but predicted

negative

● **FP** / **False Positive**:

○ when a case was negative but

predicted positive

Classification Performance Measurements Parameters 

**Classification Report**

● **Precision** – Accuracy of positive predictions.

○ Precision = TP/(TP + FP)

● **Recall**: Fraction of positives that were correctly identified.

○ Recall = TP/(TP+FN)

● **F1** Score: What percent of positive predictions were correct?

○ F1 Score = 2*\*(Recall \** Precision) / (Recall + Precision)

● **Support**: The number of actual occurrences of the class in the specified dataset.