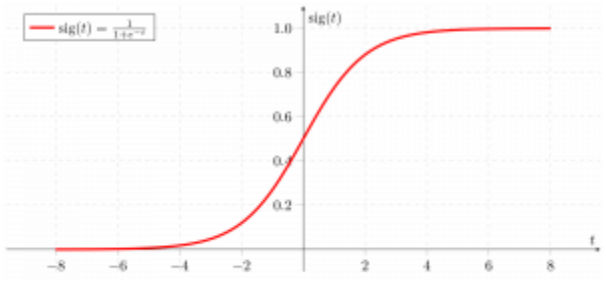
Lab5

➢ Logistic Regression:

- Logistic regression is a supervised machine learning algorithm mainly used for classification tasks where the goal is to predict the probability that an instance of belonging to a given class.
- It's referred to as regression because it takes the output of the linear regression function as input and uses a sigmoid function to estimate the probability for the given class.
- The difference between linear regression and logistic regression is that linear regression output is the continuous value that can be anything while logistic regression predicts the probability that an instance belongs to a given class or not.

➢ Sigmoid function:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$



Sigmoid function

- As shown is above fig sigmoid function converts the continuous variable data into the probability between 1 and 0.
- $\sigma(z)$ tends towards 1 as $z \to \infty$
- $\sigma(z)$ tends towards 0 as $z \to -\infty$
- $\sigma(z)$ is always bounded between 0 and 1

➢ Logistic Regression Steps:

1. Initialize weight parameters to initial values (usually small random values, rather than zero).
2. Calculate net input z as the linear combination of our feature variables x and the model weights w.

$$\mathbf{z} = w_1 x_1 + \cdots + w_m x_m = \sum_{j=1}^{m} x_j w_j$$

$$= \mathbf{w}^T \mathbf{x}$$

3. apply sigmoid activation function where $\sigma(z) = \frac{1}{1+e^{-z}}$
4. Update of the weights using gradient descent optimization method

$$\frac{\partial(J(\theta))}{\partial(\theta j)} = -\frac{1}{m} * \left(\sum_{i=1}^{m} \left[\, y^{(i)} - h_\theta\left(x^{(i)}\right)\right] * x_j^i\right)$$

$$\theta_j \leftarrow \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right) x_j^{(i)}$$

➢ K-Fold cross validation:

The general procedure is as follows:

1. Shuffle the dataset randomly.
2. Split the dataset into k groups
3. For each unique group:
   1. Take the group as a hold out or test data set
   2. Take the remaining groups as a training data set
   3. Fit a model on the training set and evaluate it on the test set
   4. Retain the evaluation score and discard the model
4. Summarize the skill of the model using the sample of model evaluation Scores.

➢ MNIST dataset:

- MNIST is short for Modified National Institute of Standards and Technology database.

- MNIST contains a collection of 70,000, 28 x 28 images of handwritten digits from 0 to 9