California State University
Los Angeles

Machine Learning

Logistic Regression

Nikhil Dhiman
Computational Molecular Biology (COMB) Lab

Agenda Overview

01 What is Machine Learning?

02 Machine Learning Settings

03 Approaches of Supervised Learning

04 Linear Regression

05 Logestic Regression

06 Sigmod Function

07 Cost Function in Linear Regression

What is Machine Learning?

Definition

- Designing and constructing algorithms or methods that give computers the ability to learn from past data, without being explicitly programmed, and then make predictions on future data.

- A set of algorithms that can automatically detect and extract patterns in past data, and then use the extracted patterns to predict future data, or to perform other kinds of decision-making.

Machine Learning Settings

Supervised learning

Learning from labeled observations

Unsupervised learning

Learning from unlabeled observations

Semi-supervised learning

Labels are provided only for a part of the training data

Reinforcement learning

Learning from an agent taking actions in an environment so as to maximize a long-term reward.

Transfer learning

Learning from a dataset while solving a problem, and then applying the extracted knowledge to a different but related dataset/problem.

Active learning

Similar to Semi-Supervised Learning, but the algorithm is able to interactively query the user or some other information source to obtain the labels as needed.

Two Important Approaches of Supervised Learning

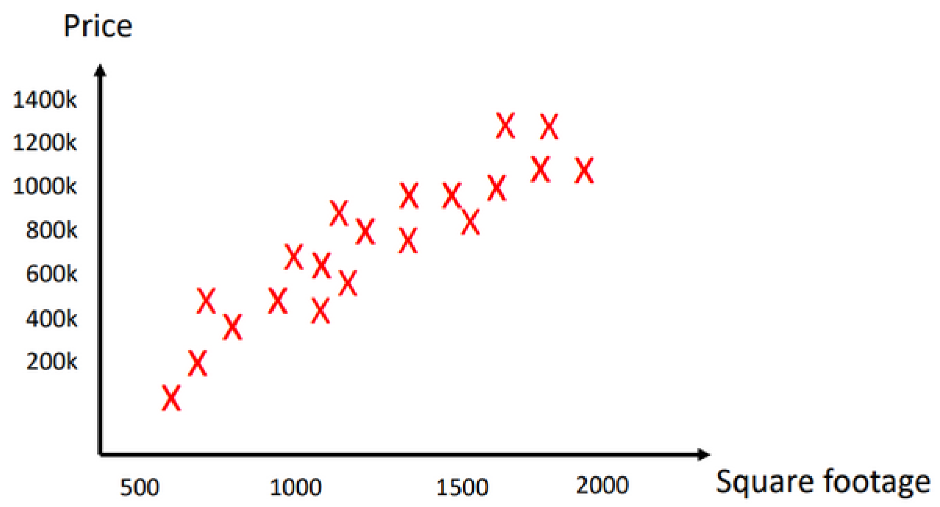
- **Classification**: Predict a discrete valued output for each observation. – Labels are discrete (categorical) – Labels can be binary (e.g., rainy/sunny, spam/non-spam,) or non-binary (e.g., rainy/sunny/cloudy, object recognition (100classes))

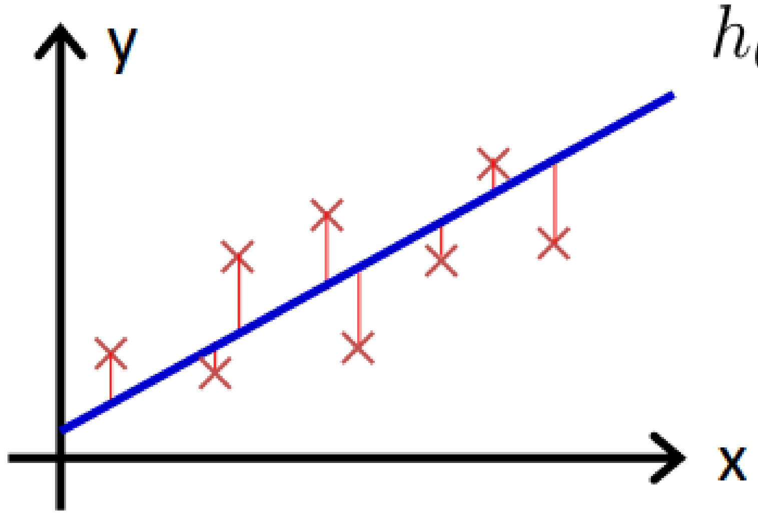
- **Regression**: Predict a continuous valued output for each observation. – Labels are continuous (numeric), e.g., stock price, housing price – Can define 'closeness' when comparing prediction with true values

Linear Regression

- A statistical method to model the relationship between a dependent variable and one or more independent variables.
- Predicts continuous outcomes.
- Simple yet powerful for many real-world applications.
- Linear Regression with single input variable (one feature). It is also called “Univariate Linear Regression.

Regression Example: Housing Price





$$h_\theta(x) = \theta_0 + \theta_1 x$$

$$\theta_0, \theta_1 = ?$$

Cost Function in Linear Regression

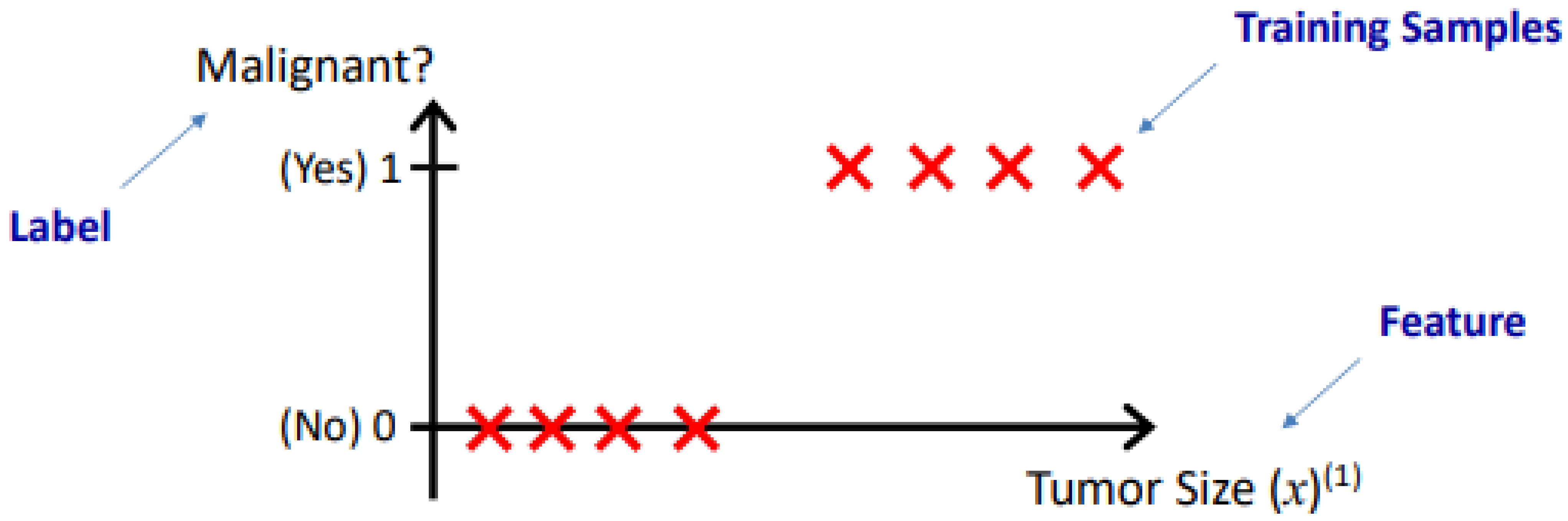
- The Cost Function, J(θ0,θ1), quantifies the error between the predicted values and actual values in the training data.
- It helps measure how well the model fits the data.

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

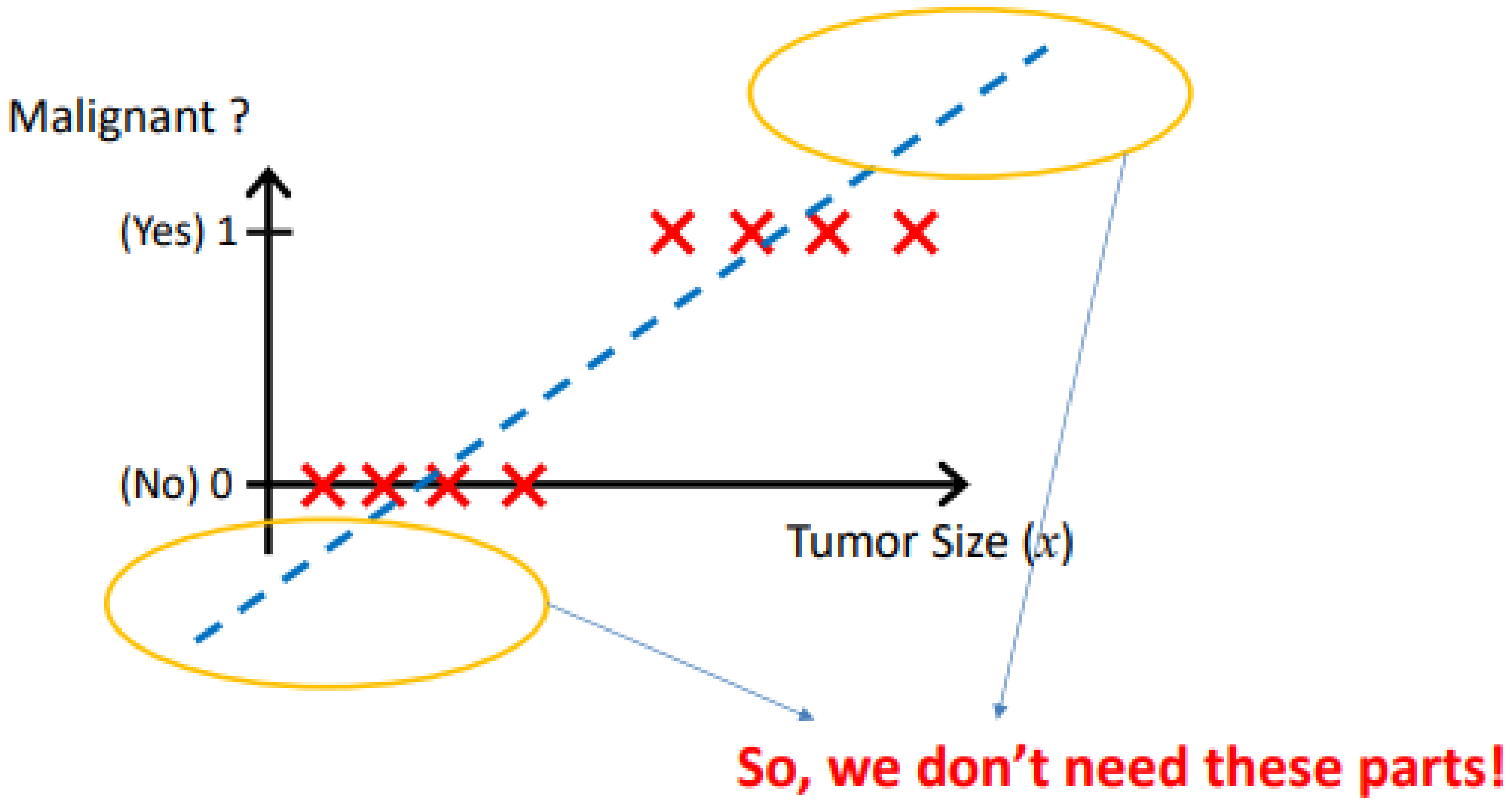
Logistic Regression

- Logistic Regression is a statistical model used for binary classification, where the goal is to predict one of two possible classes (e.g., 0 or 1) based on input features.
- It applies a sigmoid function to the linear combination of input features and weights, which maps any real-valued number to a probability between 0 and 1.
- The model is trained using maximum likelihood estimation, typically minimizing the log loss (cross-entropy) function to find the best-fitting parameters for classification.

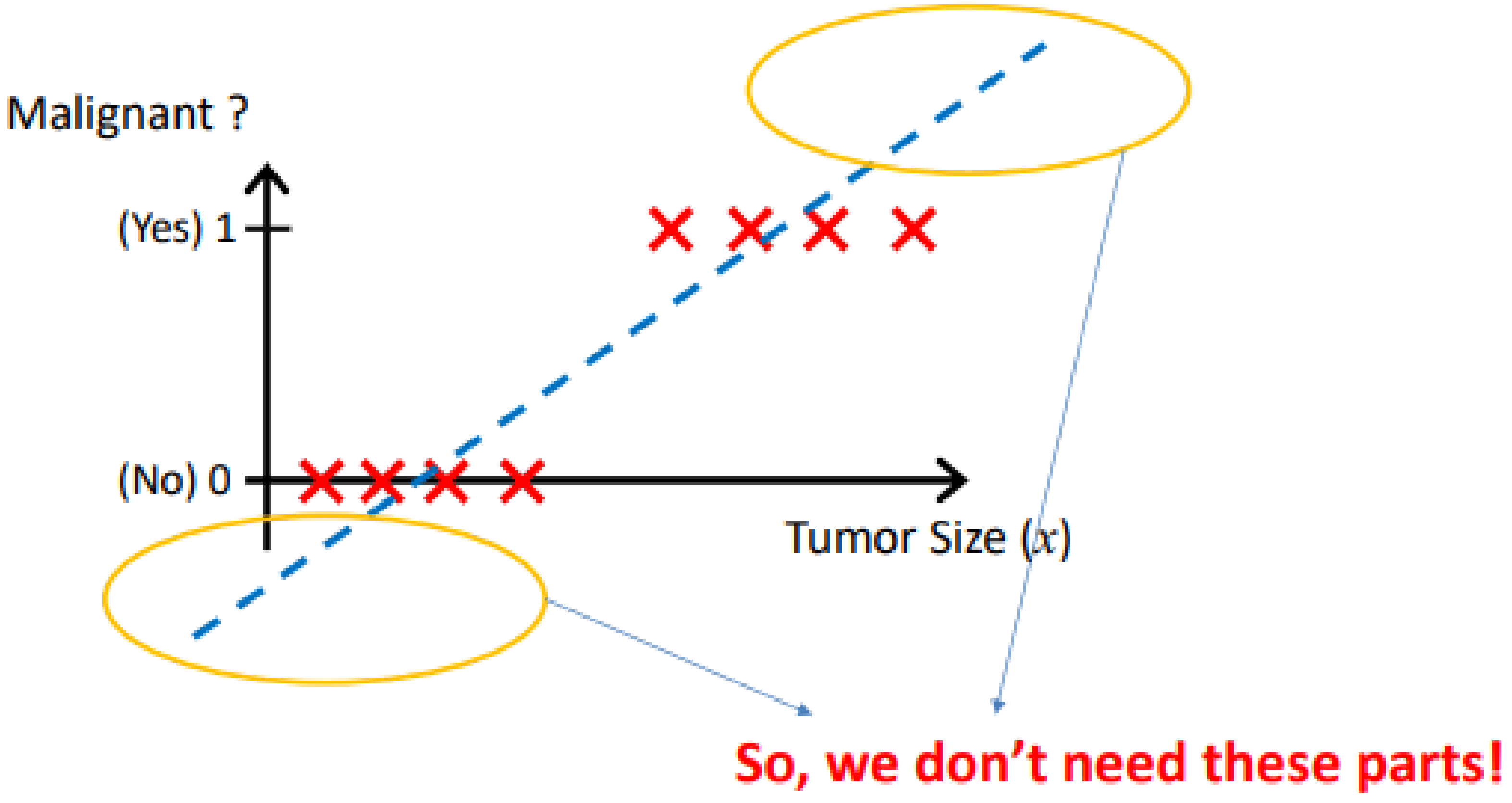
:Predicting if a Tumor is Malignant or Not based on tumor size (so, we need a classifier):



we know that for our classifier, the output should be either 1 or 0!



we know that for our classifier, the output should be either 1 or 0!

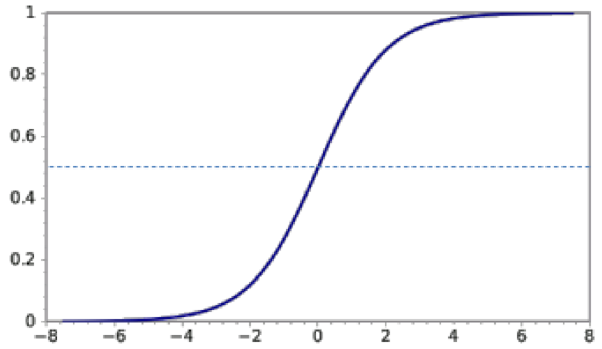


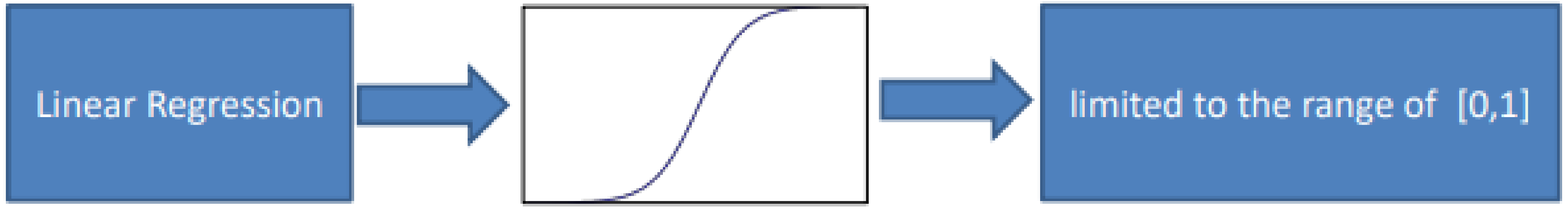
$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$

Linear Regression Model $h\theta(x) = \theta^T x$

Sigmoid Function

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





$$h_\theta(\boldsymbol{x}) = \boldsymbol{\theta}^T \boldsymbol{x}$$

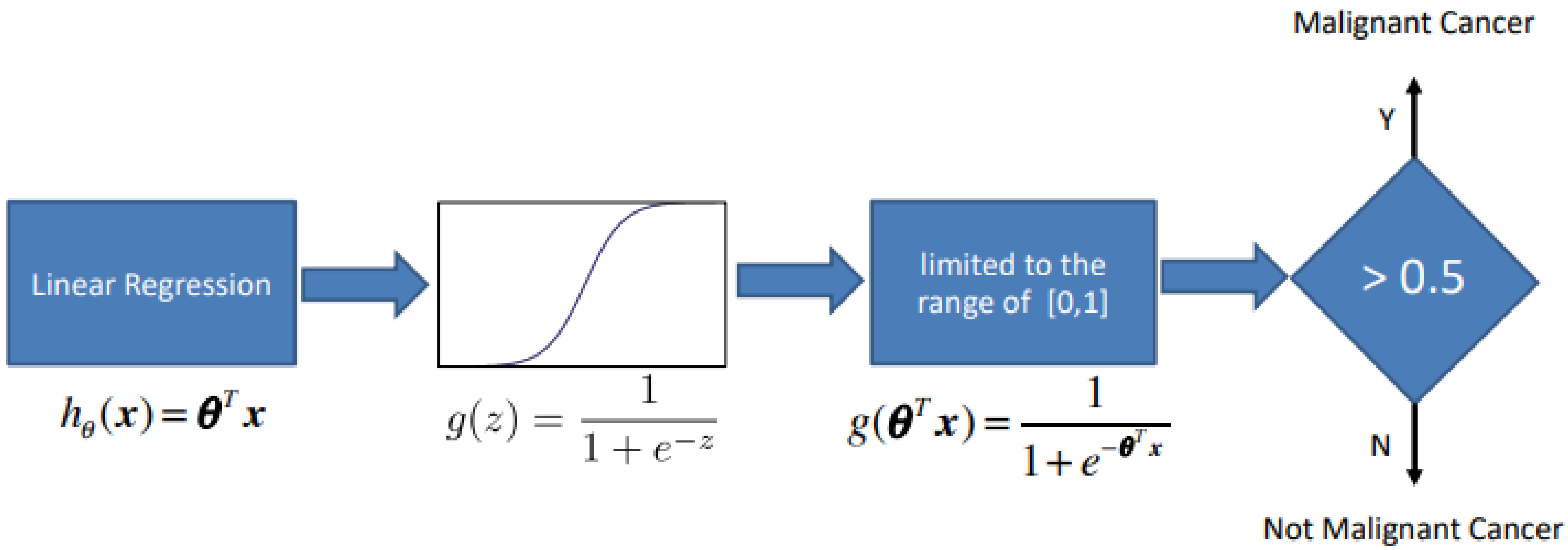
$$g(z) = \frac{1}{1 + e^{-z}}$$

$$g(\boldsymbol{\theta}^T \boldsymbol{x}) = \frac{1}{1 + e^{-\boldsymbol{\theta}^T \boldsymbol{x}}}$$

New approach for output prediction:

$$h_\theta(\boldsymbol{x}) = g(\boldsymbol{\theta}^T \boldsymbol{x}) = \frac{1}{1 + e^{-\boldsymbol{\theta}^T \boldsymbol{x}}}$$

So, Now the NEW $h_\theta(x)$ is limited to the range of [0,1].



New Cost function:

$$J(\boldsymbol{\theta}) = J(\theta_0, \theta_1, \dots, \theta_n)$$

$$= -\frac{1}{m} \sum_{i=1}^{m} \left[y^{(i)} \log h_\theta(\boldsymbol{x}^{(i)}) + (1 - y^{(i)}) \log(1 - h_\theta(\boldsymbol{x}^{(i)})) \right]$$

Note: $y = 0$ or 1 always

California State University
Los Angeles

Thank You

Computational Molecular Biology (COMB) Lab