LECTURER: TAI LE QUY

MACHINE LEARNING – SUPERVISED LEARNNG

UNIT 2

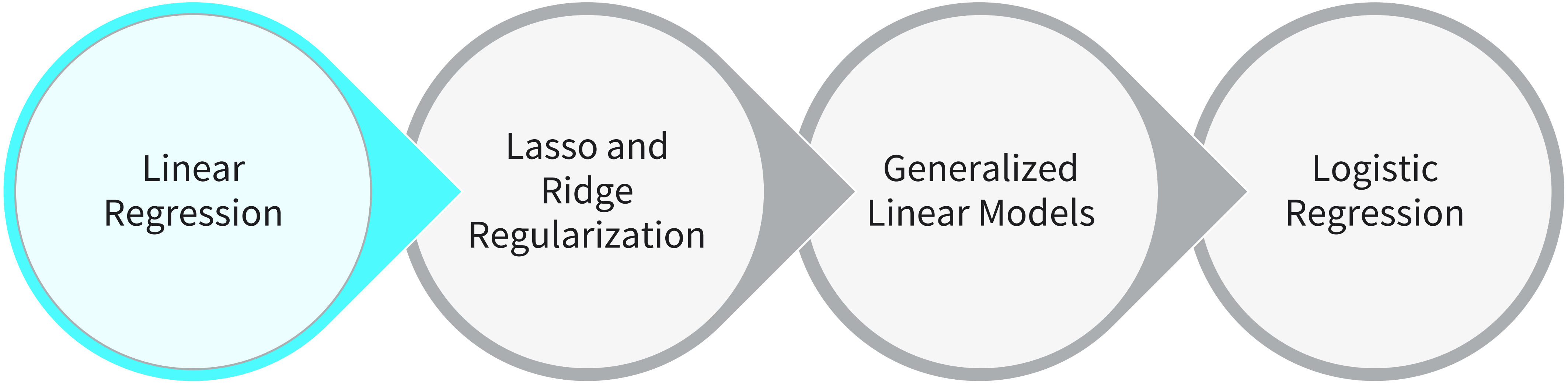
REGRESSION

STUDY GOALS

- understand the concept of regression and when to use it.
- evaluate a regression model’s performance.
- utilize regularization techniques and understand where they are implemented.
- apply different well-known regression models with the use of Python.

1. How do regression models work, and what is the math behind them?
2. How can we evaluate the performance of a regression model?
3. How can we reduce bias and variance in regression models?
4. How can we apply a regression model with Python?

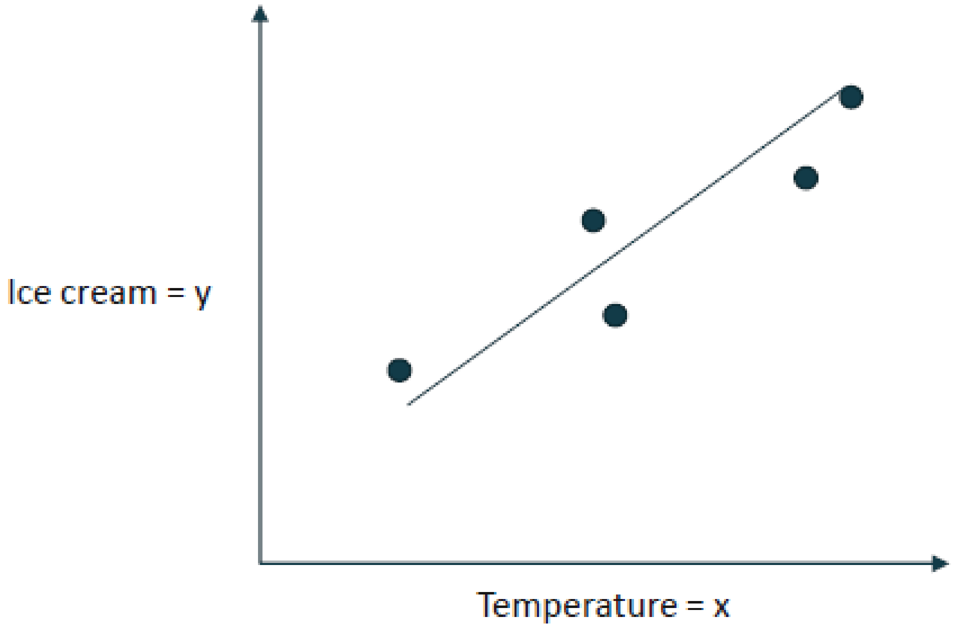
REGRESSION


Linear Regression
Lasso and Ridge Regularization
Generalized Linear Models
Logistic Regression

Important Regression Terms and Their Synonyms	
Term	Synonyms
Target variable ($\mathbf{Y}$)	Label, y variable, dependent variable
Input ($\mathbf{X}$)	X variables, independent variables, pre-dictors, features
Coefficient (ω_i)	Weight, slope, regression coefficient
(y-axis) Intercept (ω_0)	Bias
Loss function	Cost function, target function, objective function, error function

Source of image: Course book

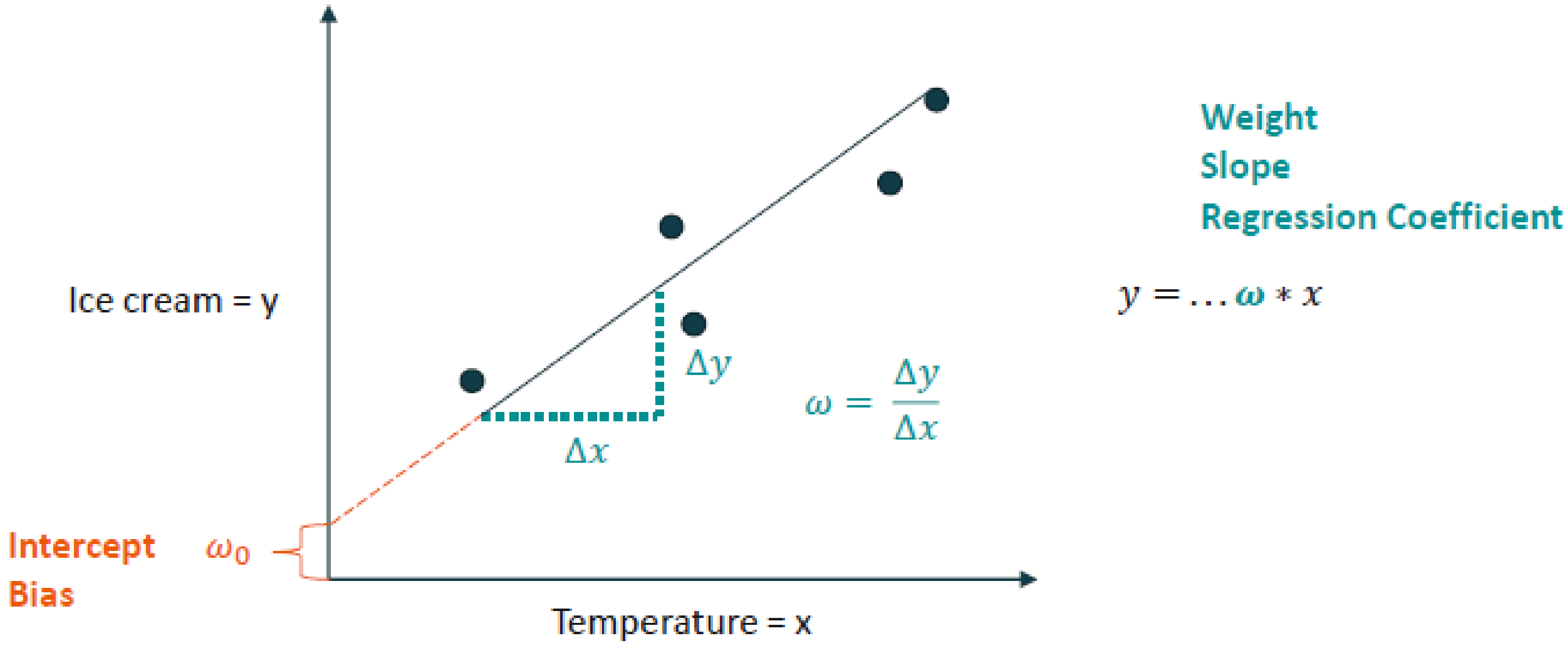
LINEAR REGRESSION



$$y = ? * x$$

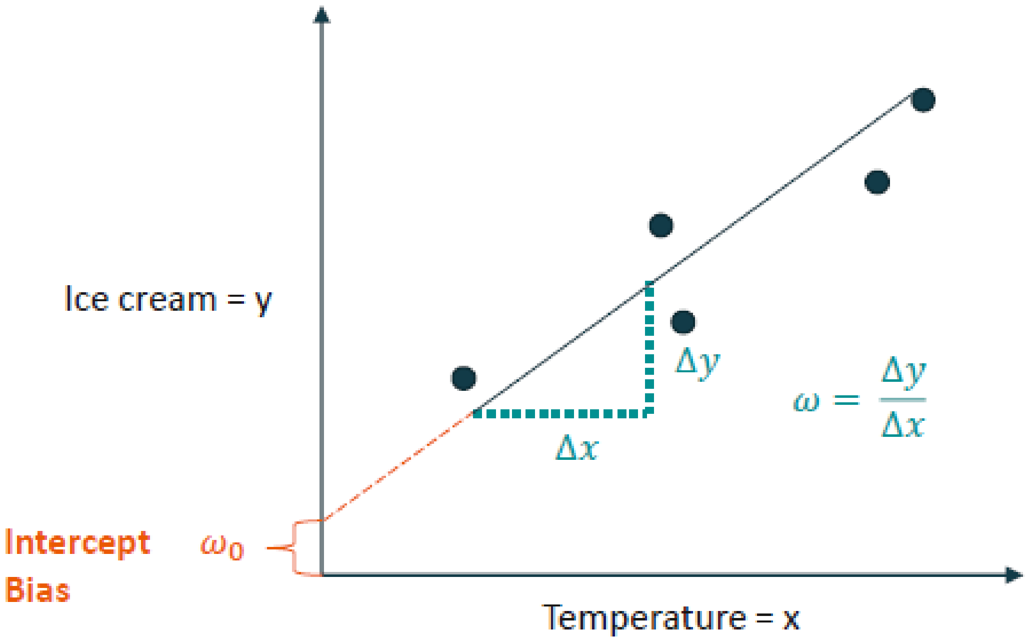
Source of image: https://learn.iu.org/

LINEAR REGRESSION



Source of image: https://learn.iu.org/

LINEAR REGRESSION

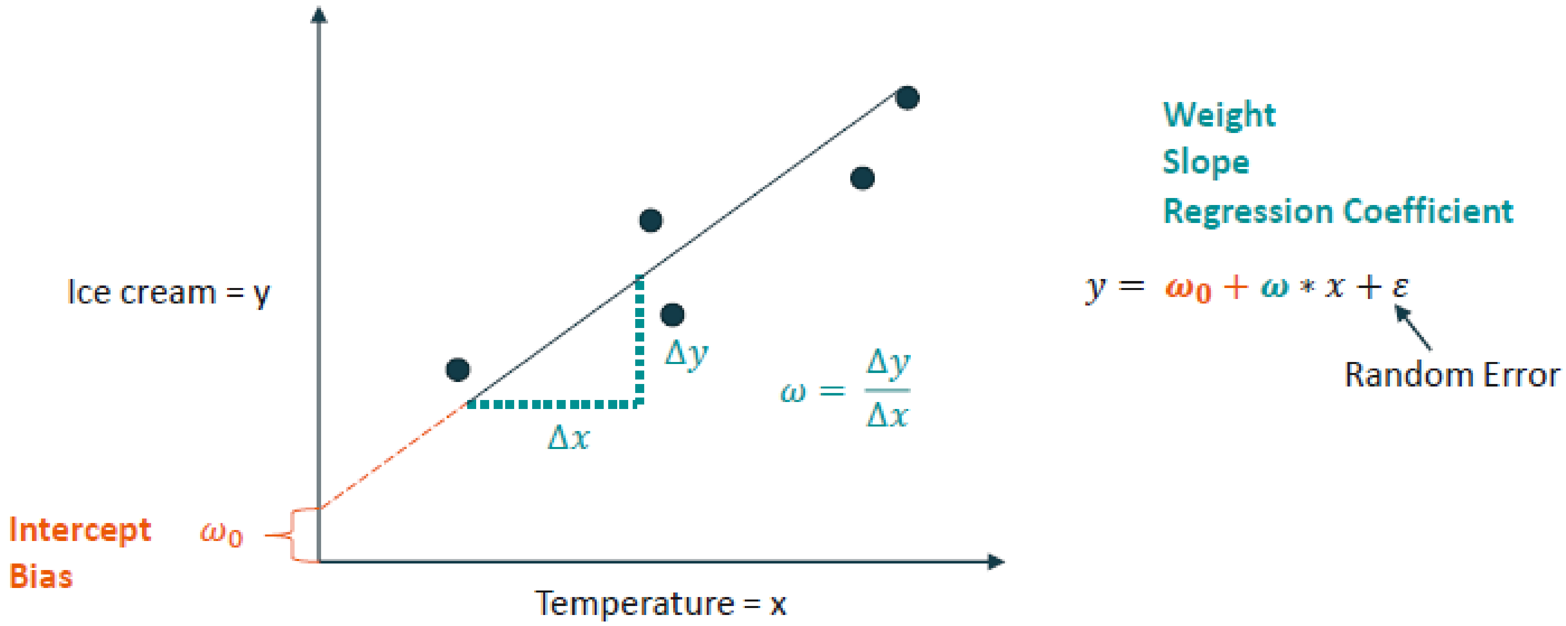


Weight
Slope
Regression Coefficient

$$y = \omega_0 + \omega * x$$

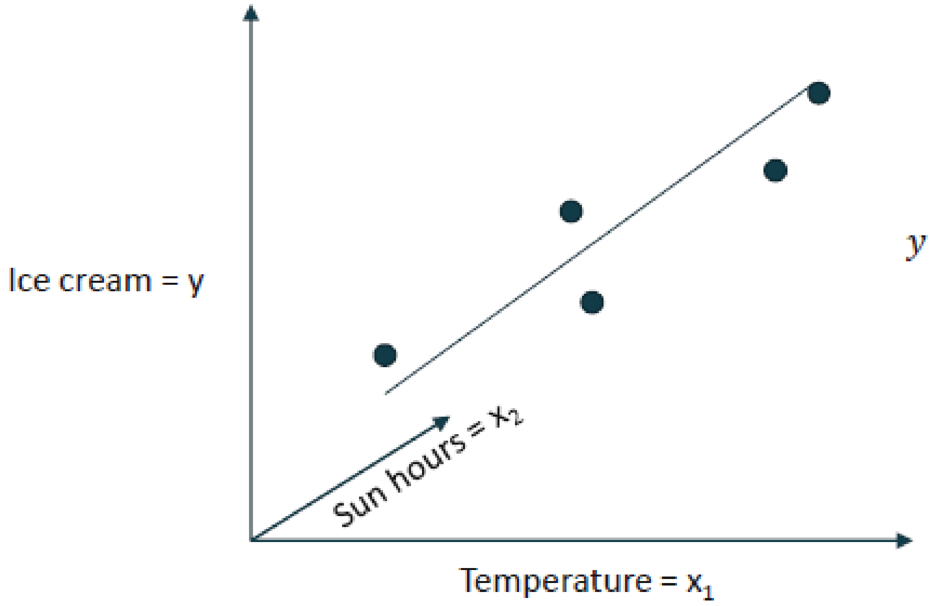
Source of image: https://learn.iu.org/

LINEAR REGRESSION



Source of image: https://learn.iu.org/

LINEAR REGRESSION

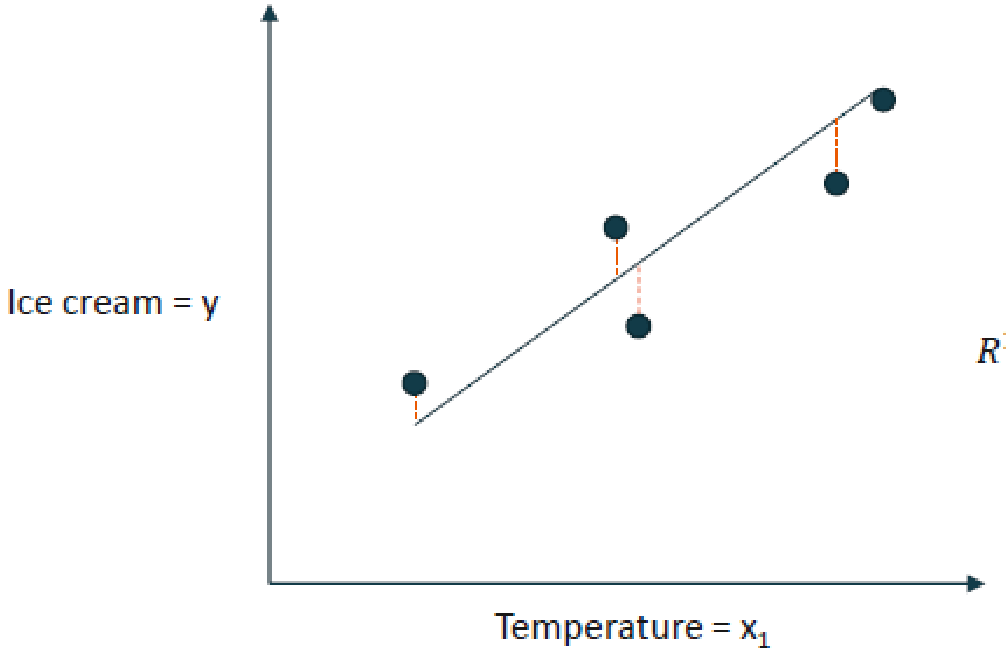


$$y = \omega_0 + \omega_1 * x_1 + \omega_2 * x_2 + \varepsilon$$

$$y = \omega_0 + \sum_{i=1}^{n} \omega_i * x_i + \varepsilon$$

Source of image: https://learn.iu.org/

LINEAR REGRESSION



$$Residual = R_i = y_i - \hat{y}_i$$

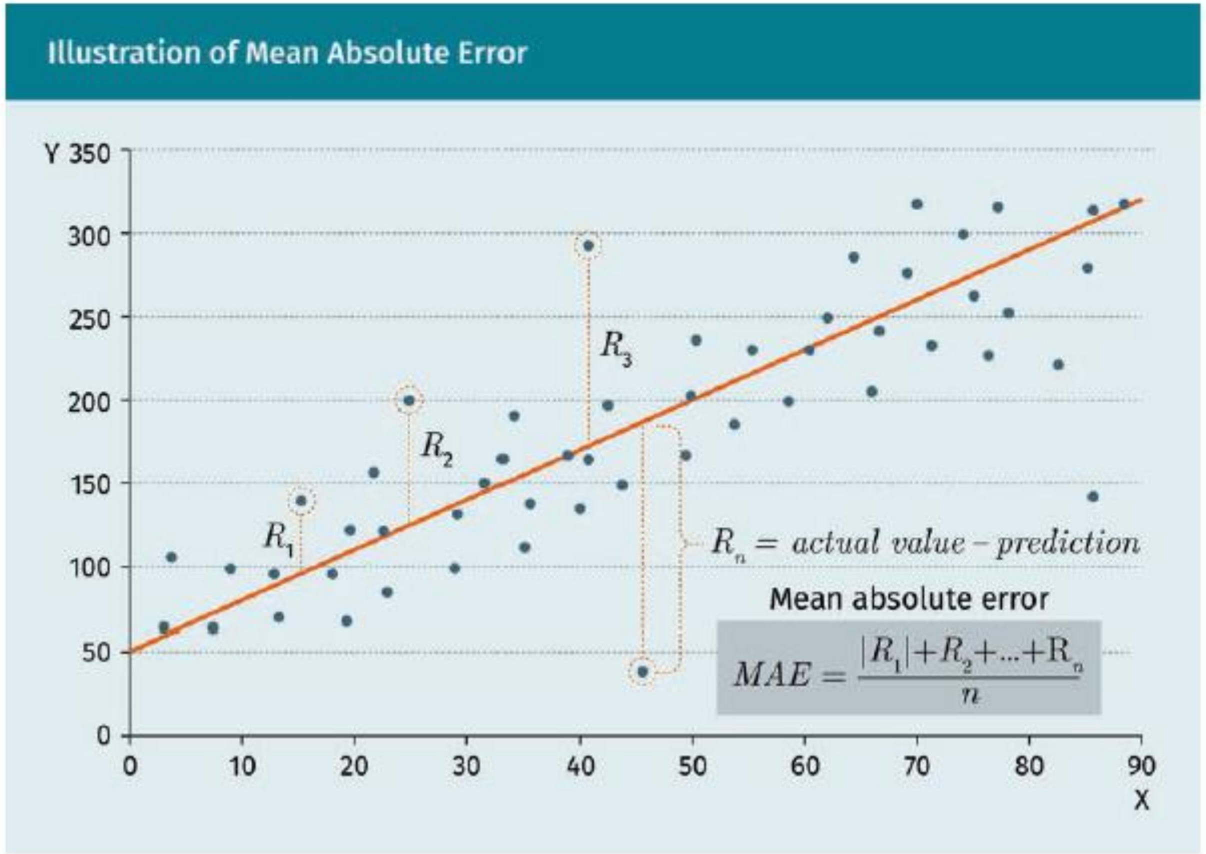
$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$R^2 = 1 - \frac{RSS}{Variance\ in\ the\ data}$$

$$R^2 = 1 - \frac{RSS}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}$$

Source of image: https://learn.iu.org/

Mean absolute error (MAE): absolute difference between all predictions and the actual values

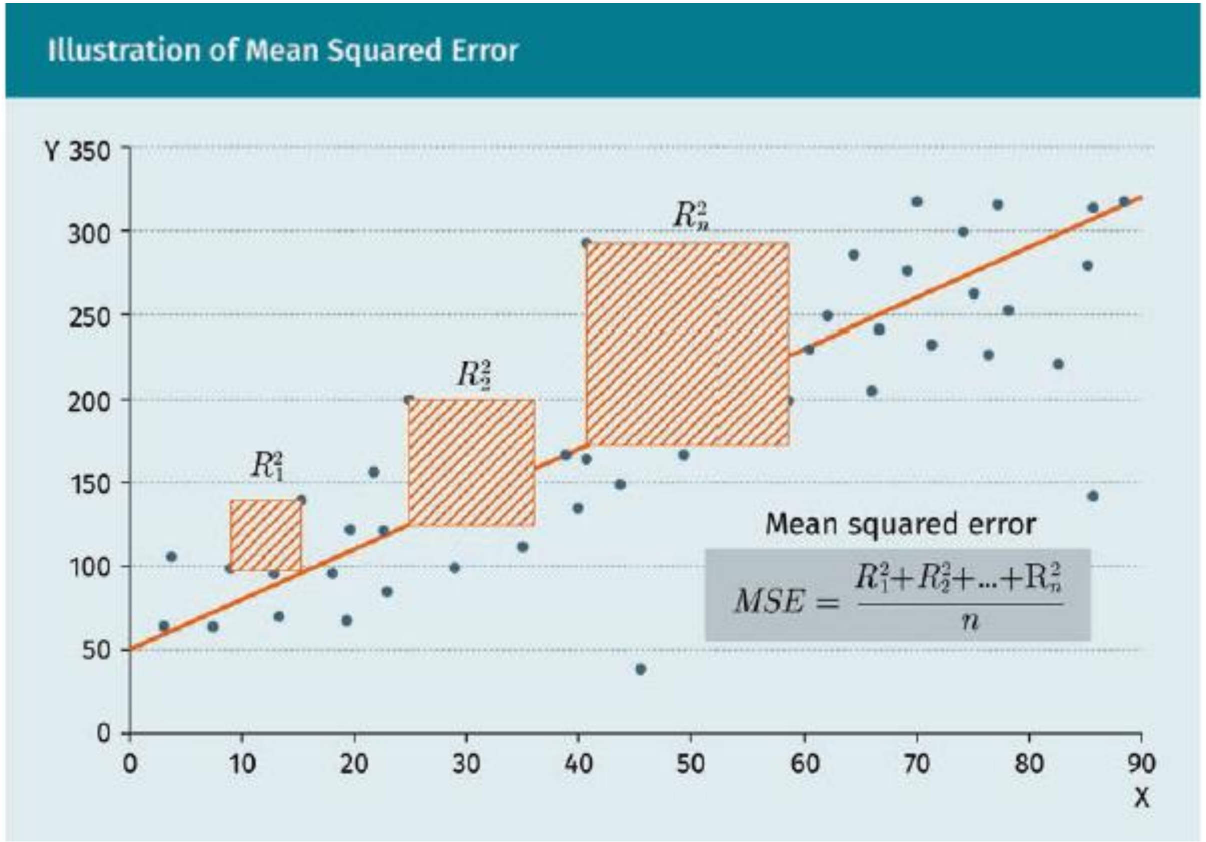


$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$$

Source of image: Course book

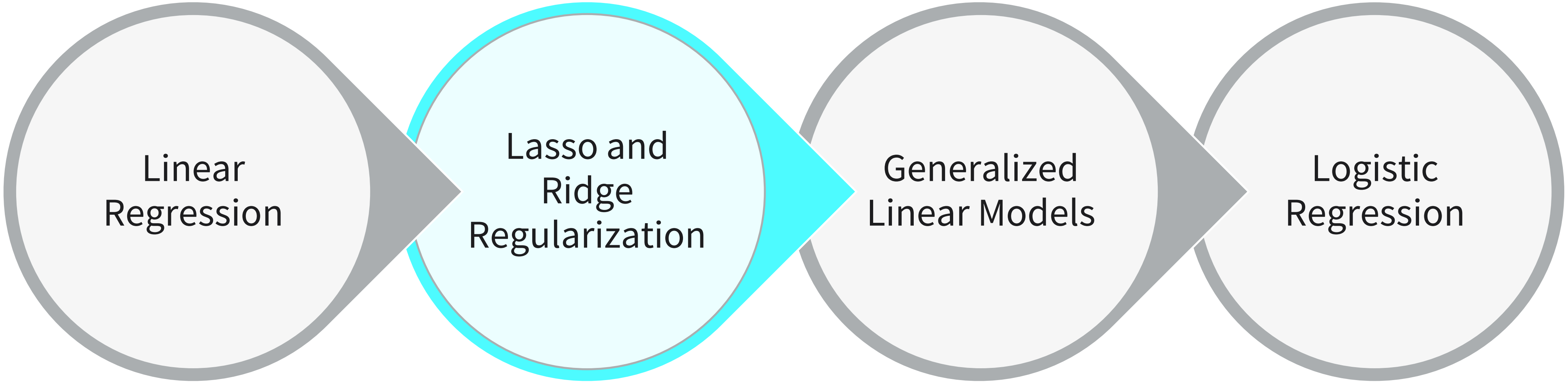
METRIC FOR MEASURING THE PREDICTION PERFORMANCE OF A REGRESSION MODEL

- Mean squared error (MSE): The deviations between the actual and the predicted values are taken to the square.
- Root mean square error (RMSE): the result is telegraphed in the unit of the label being predicted

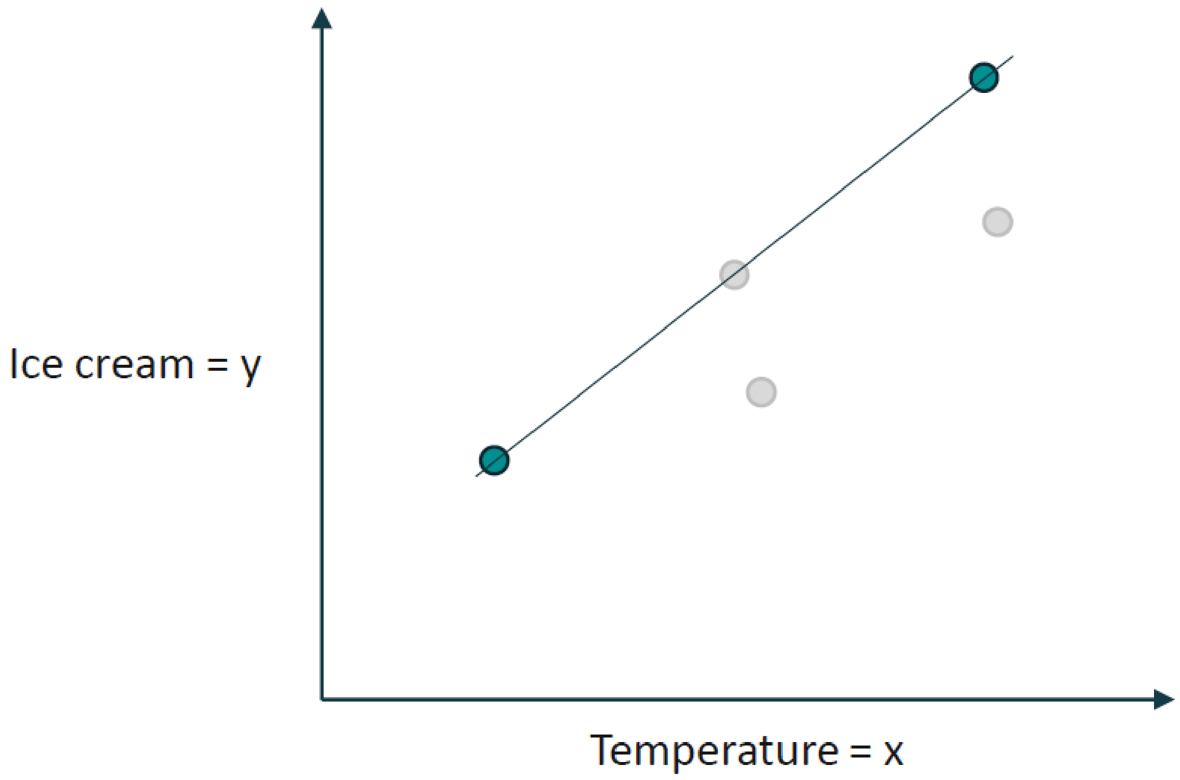


$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y}_i\right)^2 \qquad RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y}_i\right)^2}$$

Source of image: Course book


Linear Regression
Lasso and Ridge Regularization
Generalized Linear Models
Logistic Regression

LASSO AND RIDGE REGULARIZATION

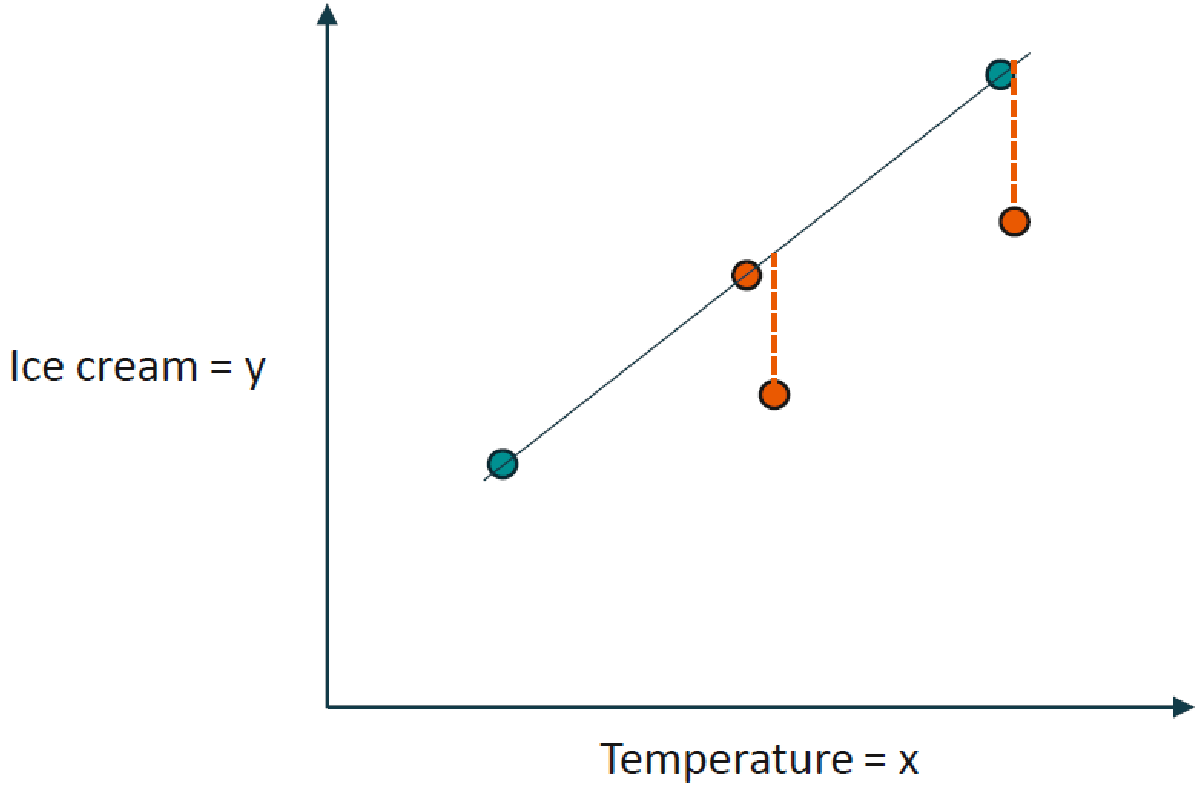


$$y = \omega_0 + \sum_{i=1}^{n} \omega_i x_i + \varepsilon$$

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Source of image: https://learn.iu.org/

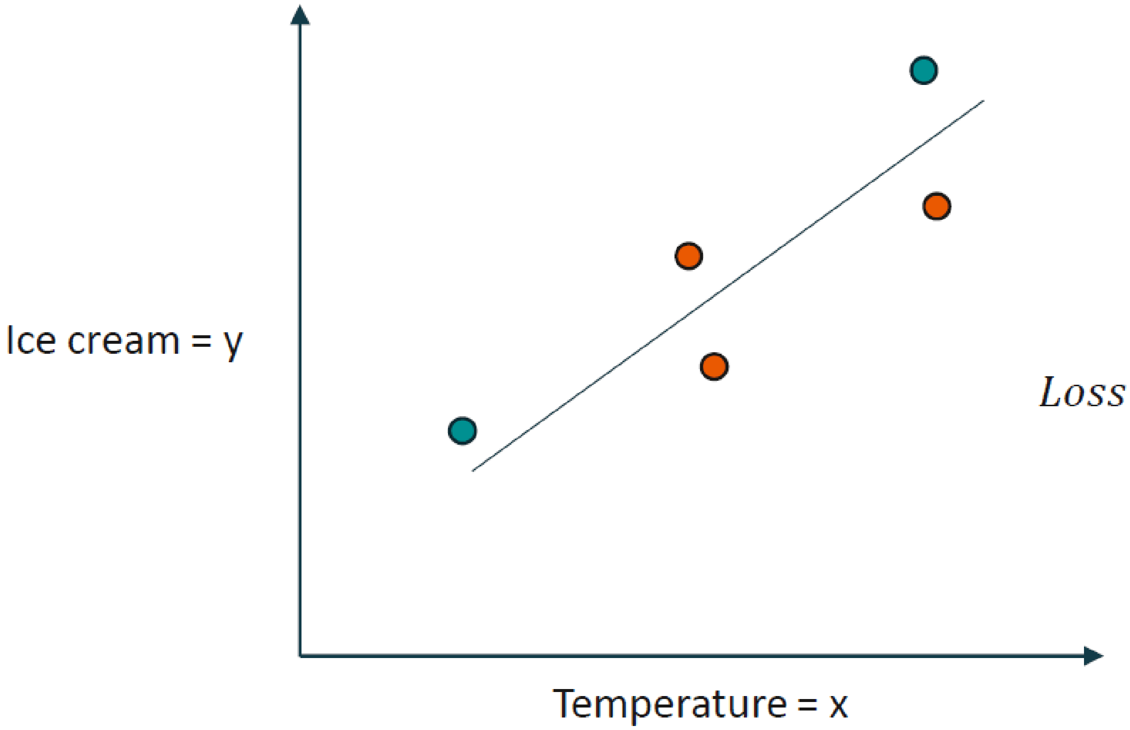
LASSO AND RIDGE REGULARIZATION



$$y = \omega_0 + \sum_{i=1}^{n} \omega_i x_i + \varepsilon$$

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Source of image: https://learn.iu.org/



$$y = \omega_0 + \sum_{i=1}^{n} \omega_i x_i + \varepsilon$$

$$Loss = \underbrace{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}_{\textbf{small}} + \underbrace{\boldsymbol{\alpha} \sum_{i=1}^{n} \boldsymbol{\omega}_i^2}_{\textbf{small}}$$

The penalty size, also known as the L2 **norm**

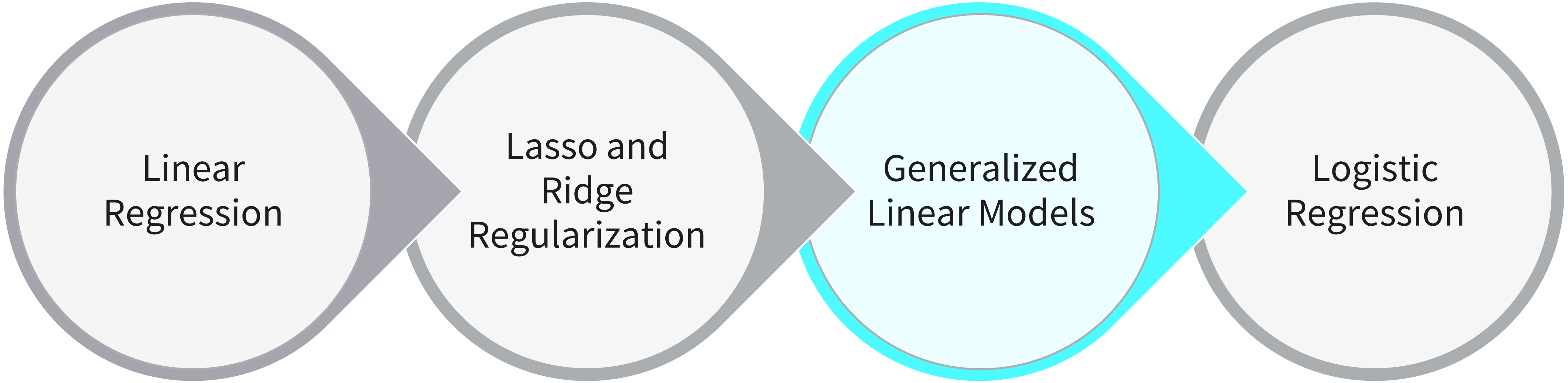
Source of image: https://learn.iu.org/

- Lasso (least absolute shrinkage and selection operator)
  - It differs from ridge regression only in that the L2 norm is exchanged for the L1 norm
- The loss function:

$$RSS + \alpha \sum_{i=1}^{n} |\omega_i|$$

- Elastic net: selects important coefficients, as does lasso regression, and is effective in handling correlated features, as is ridge regression.
  - r=0: lasso regression
  - r=1: ridge regression

$$\alpha \sum_{i=1}^{n} \left(r\omega_i^2 + (1-r)|\omega_i| \right)$$

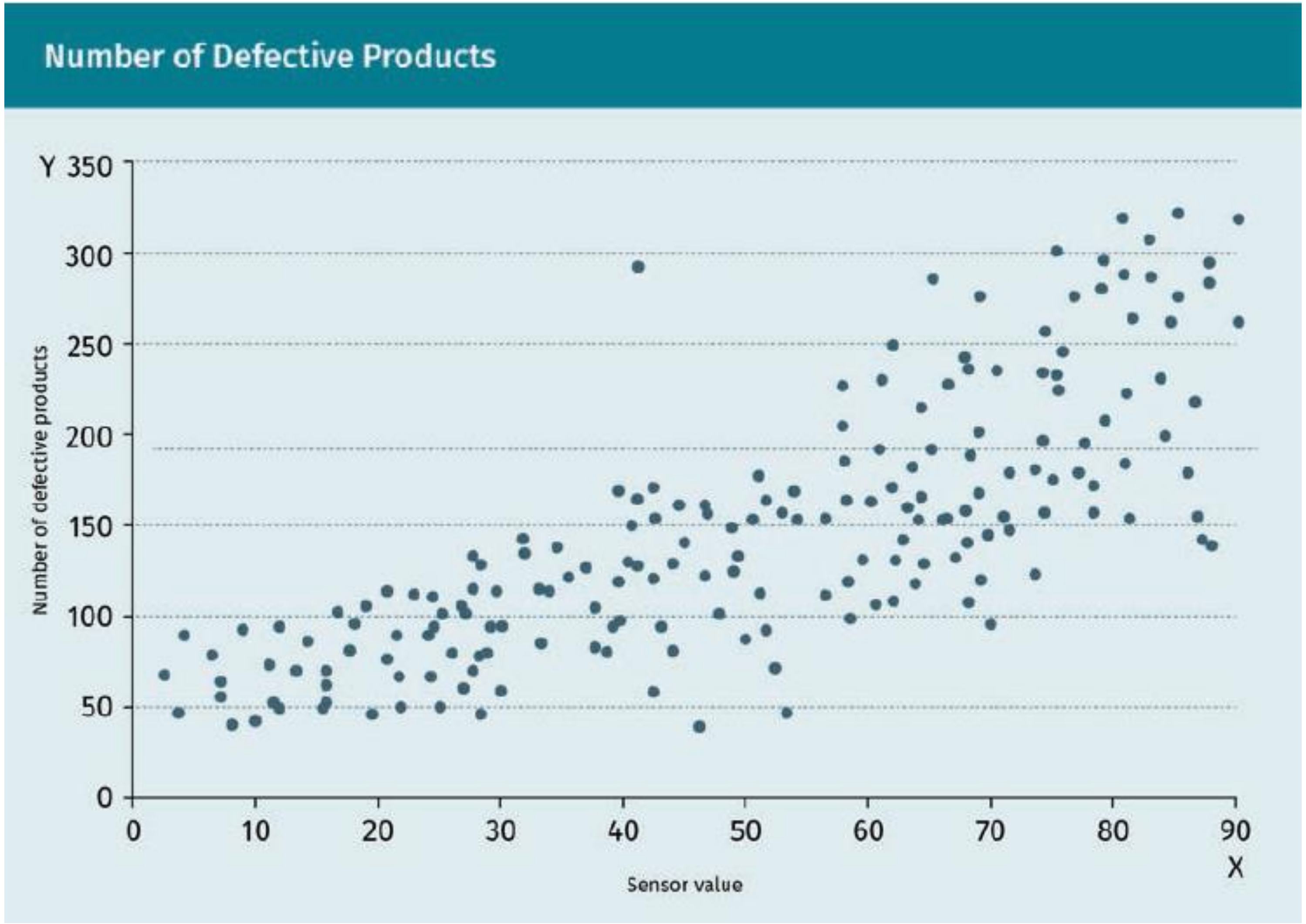

Linear Regression
Lasso and Ridge Regularization
Generalized Linear Models
Logistic Regression

- Generalized linear models (GLMs) are a category of expanded linear regression models.
- General models are developed by relaxing the assumptions of linear models.
- GLMs all essentially comprise the following three components:

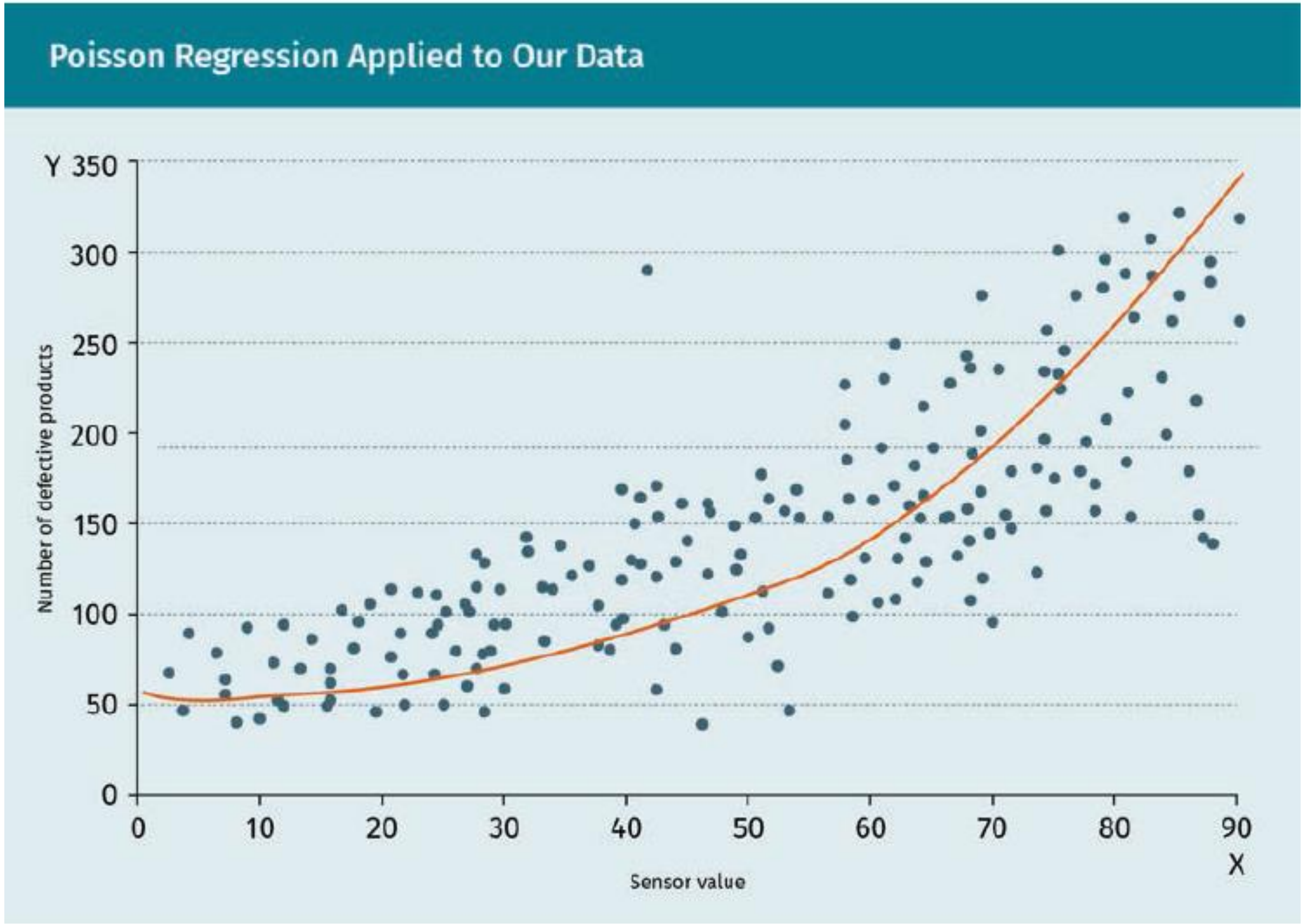
1. A linear predictor $\eta_i = \omega_0 + \omega_1 x_{1i} + \ldots + \omega_p x_{pi}$
2. A probability distribution that generates the target variable $\mathbf{Y}$
3. A monotone differentiable **link function** $g(\mu_i) = \eta_i$ describing how the mean depends on the linear predictor $\eta_{i.}$

Some Probability Distributions and Their Link Functions

Distribution	Use	Notation	Link function
Gaussian	Linear repose	$N(\mu, \sigma 2)$	“Identity”: μ
Poisson	Counts of events	$N(\mu)$	$Log(\mu)$
Bernoulli	Outcome of sin-gle yes/no occurren-ces	$Bern(p)$	$Logit(\mu)$
Binomial	Count of yes occurrences out of n yes/no events	$Bin(n, \mu)/n$	$Logit(\mu)$

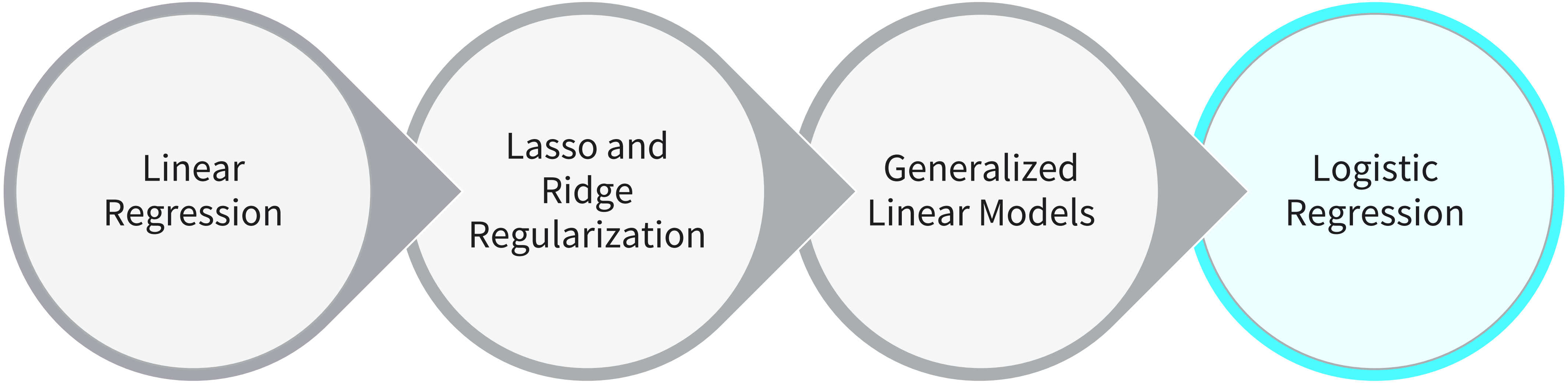


Source of image: Course book

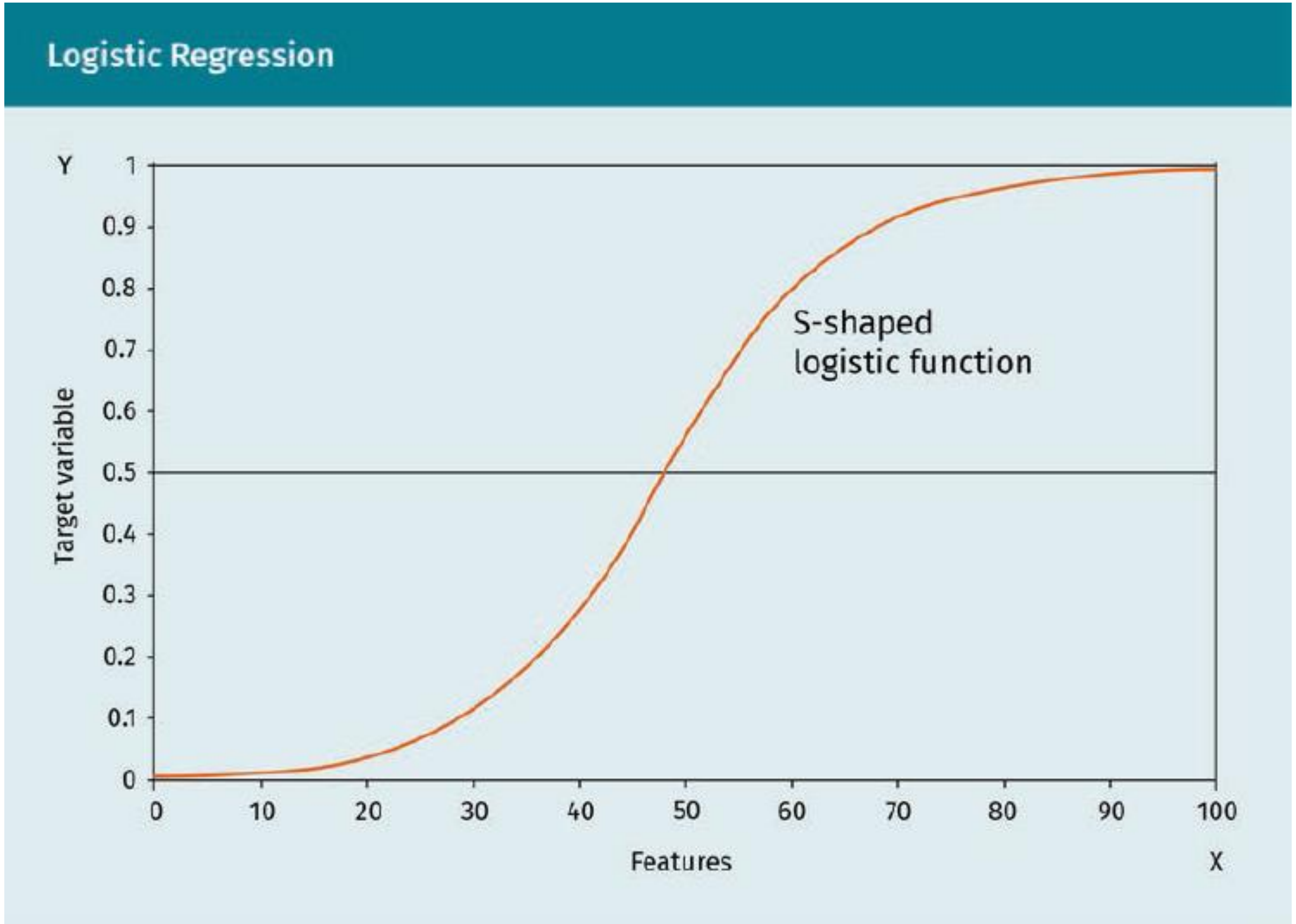


Poisson regression

$$ln\ \eta_i = \omega_0 + \omega_1 x_{1i}$$
$$y_i \sim Poisson(\eta_i)$$

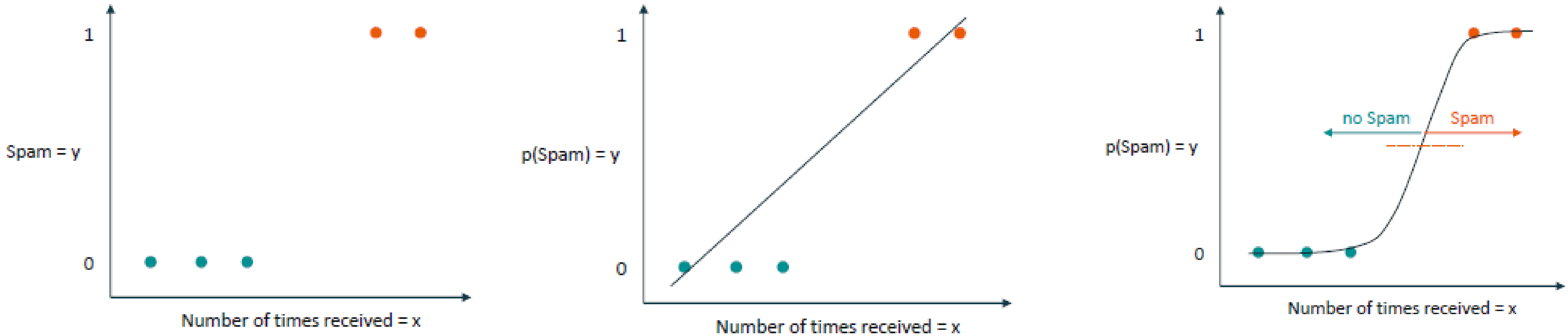

Linear Regression
Lasso and Ridge Regularization
Generalized Linear Models
Logistic Regression

LOGISTIC REGRESSION



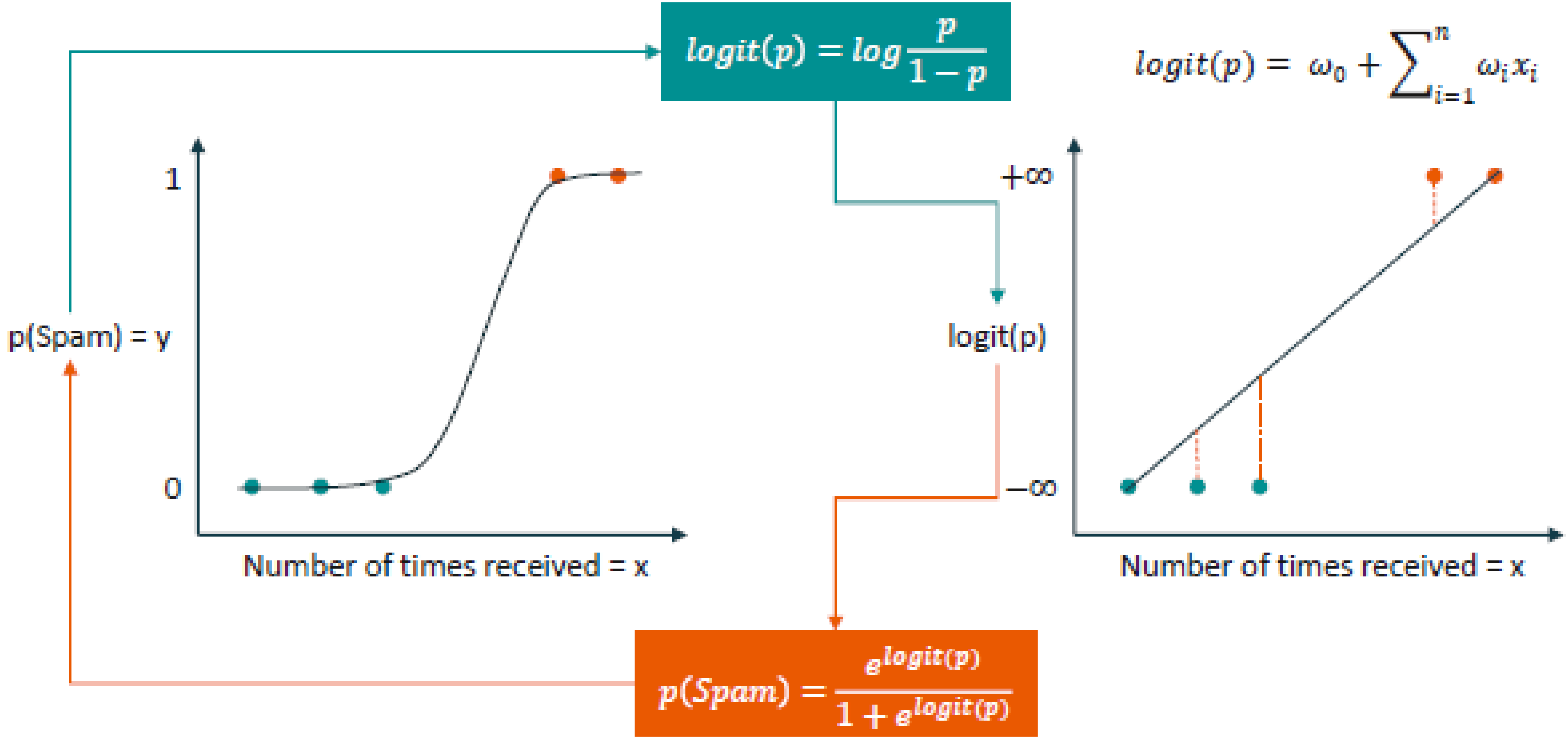
Source of image: Course book

LOGISTIC REGRESSION

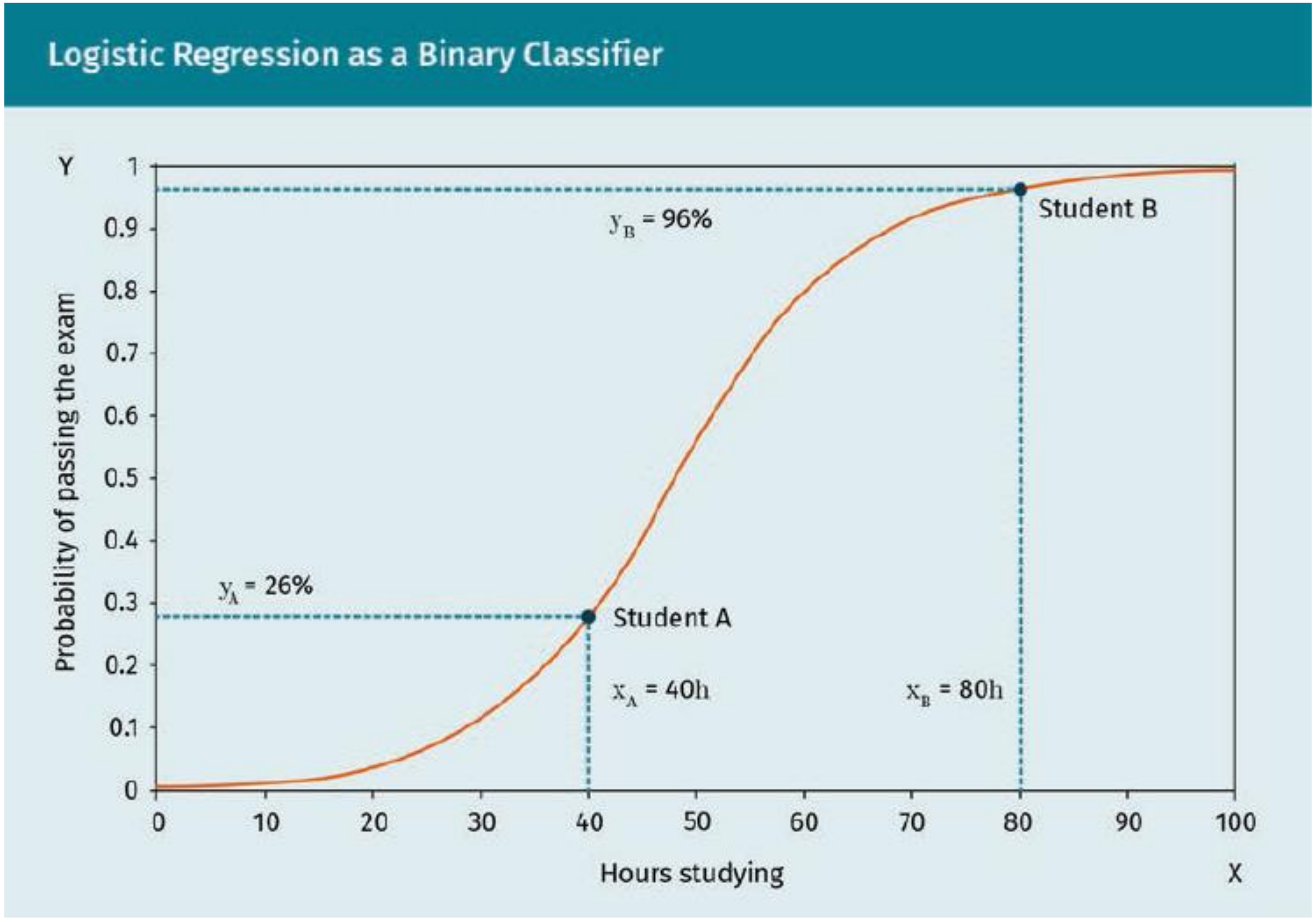


Source of image: https://learn.iu.org/

LOGISTIC REGRESSION



Source of image: https://learn.iu.org/



Logistic function

$$p\left(X\right) = \frac{e^{\omega_0 + \omega_1 X}}{1 + e^{\omega_0 + \omega_1 X}}$$

Multiple features

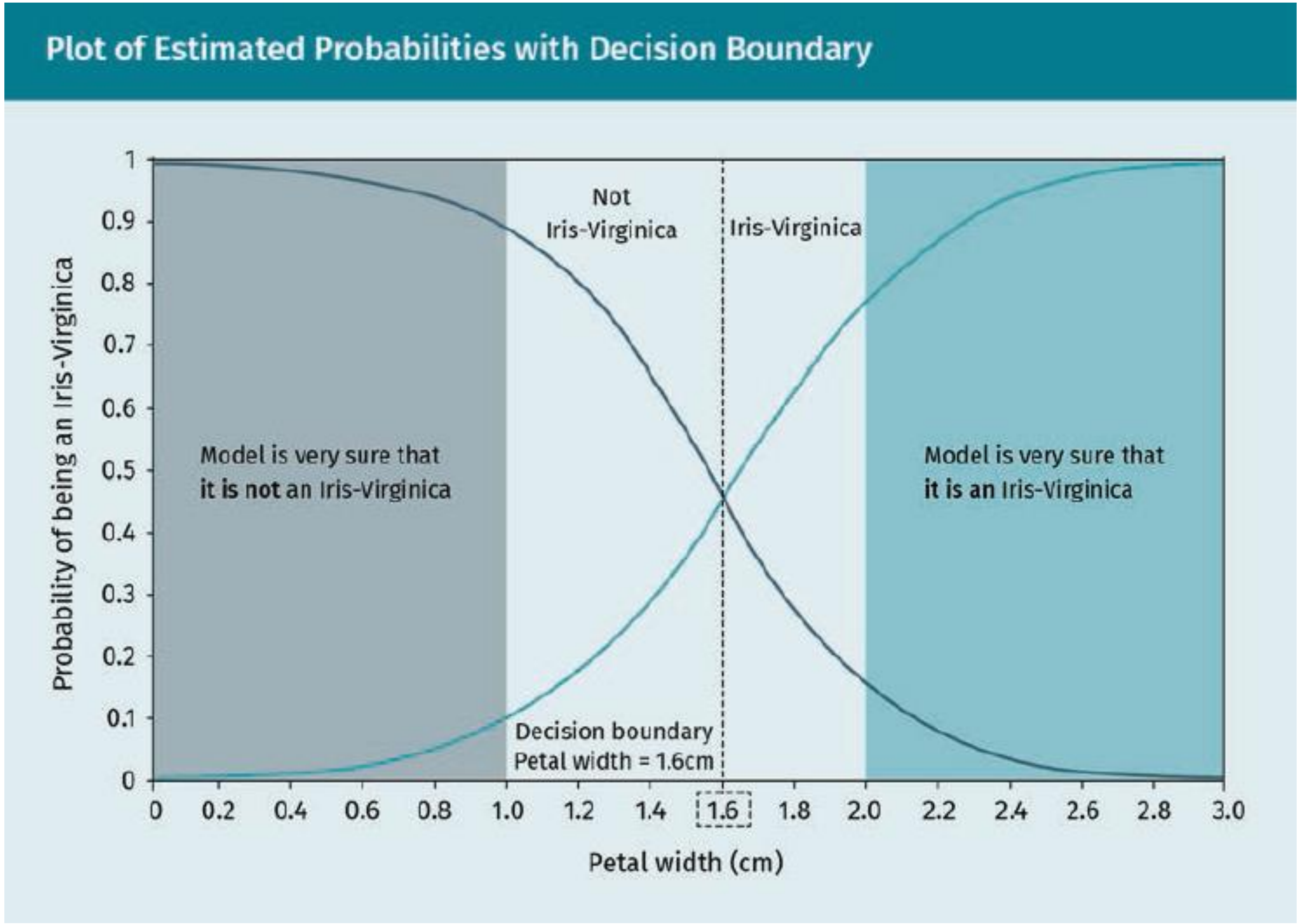
$$p\left(X\right) = \frac{e^{\omega_0 + \omega_1 X_1 + \ldots + \omega_p X_p}}{1 + e^{\omega_0 + \omega_1 X_1 + \ldots + \omega_p X_p}}$$

Source of image: Course book

- Model Training via Maximum Likelihood
  - Estimates for our regression coefficients **ω** such that the predicted probability **p(X)** for each observation matches the observed output value Y of the observation as closely as possible

$$l(\omega_0, \omega_1) = \prod_{i:\ y_i = 1} p(X_i) \prod_{i':\ y'_i = 0} [1 - p(x'_i)]$$

LOGISTIC REGRESSION – EXAMPLE



Source of image: Course book

STUDY GOALS REVIEW

- understand the concept of regression and when to use it.
- evaluate a regression model’s performance.
- utilize regularization techniques and understand where they are implemented.
- apply different well-known regression models with the use of Python.

UNIT 2

TRANSFER TASK

Stock market prediction: Case study

- Stock market prediction is the act of trying to determine the future value of a company stock or other financial instrument traded on an exchange. The successful prediction of a **stock's future price** could yield significant profit.

- Create a rough project plan to achieve this goal. For each phase of this plan, explain how regression techniques might be applied

Please present your results.
The results will be discussed in plenary.