Mr. Touch Sopheak, Mr. Phok Ponna, Mr. Pen Chentra
Department of Applied Mathematics and Statistics
Institute of Technology of Cambodia, ITC

March 11, 2025

LOGISTIC REGRESSION — TP 3: Logistic Regression

Remarks: All models here should be built from scratch.

1. Problem 1

Given the dataset "data_3_1_1.csv", build a logistic regression model with the hypothesis $h_1(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$.

2. Problem 2

Given the dataset "data_3_1_2.csv", build two logistic regression models:

- $h_1(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$

- $h_2(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2)$

Compare these two models. Which one is better?

3. Problem 3

Given the dataset "data_3_2.csv", build a multiclass (4 class) classification model to fit this data. You can choose your own hypothesis.

4. Problem 4: Mini Project no. 2

With your team, collect your own dataset which consists at least two features and one target variable. Build a logistic regression model to predict the target variable. Split your dataset into training set (80 %) and testing set (20 %). Present your model to the class next week!

Submitted by Mr. Touch Sopheak, Mr. Phok Ponna, Mr. Pen Chentra on 11 de marzo de 2025.