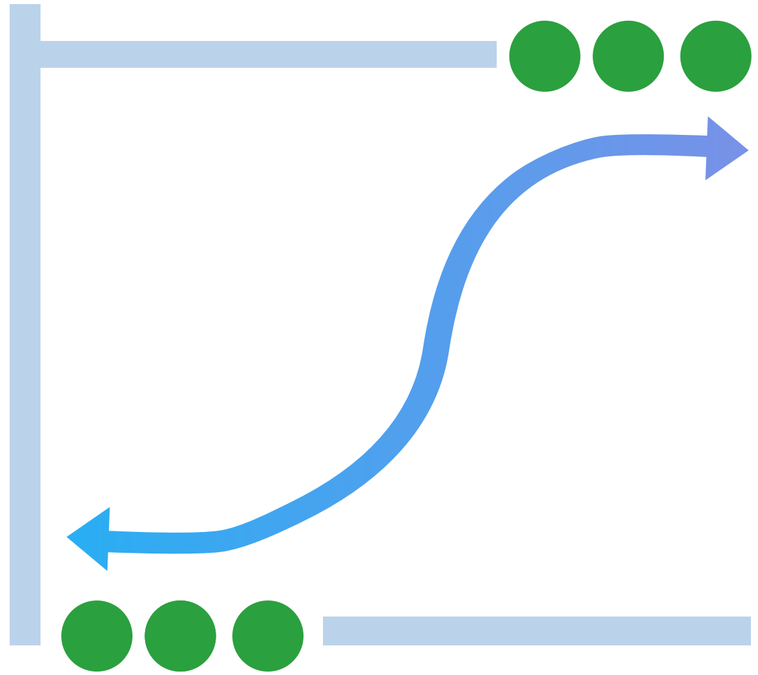
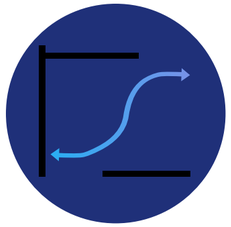
Logistics Regression



LOGISTICS REGRESSION



In this section we'll cover the a technique that predict a categorical dependent variable based on one or more independent variables

TOPICS WE'LL COVER:

- **Logistics Regression**
- **Goal of Logistics**
- **Evaluation**

GOALS FOR THIS SECTION:

- Use logistic regression for binary classification
- Implement logistic regression for binary classification
- Address overfitting using regularization, to improve model performance

LOGISTICS REGRESSION

ទីវិទ្យាស្ថានសាន់រ៉ាយស៍
SUNRISE INSTITUTE

Logistics Regression

Goals of Logistics

Evaluation

Logistic Regression is a **classification technique** used to predict the probability of a binary (true/false) outcome.

- In its simplest form, logistic regression forms an **S-shaped curve between 0 -1**, which represents the probability of a TRUE outcome for any given value of X
- The **likelihood function** measures how accurately a model predicts outcomes, and is used to optimize the "shape" of the curve
- Although it has the word *"regression"*in its name, logistic regression is not used for predicting numeric variables

Example use cases:

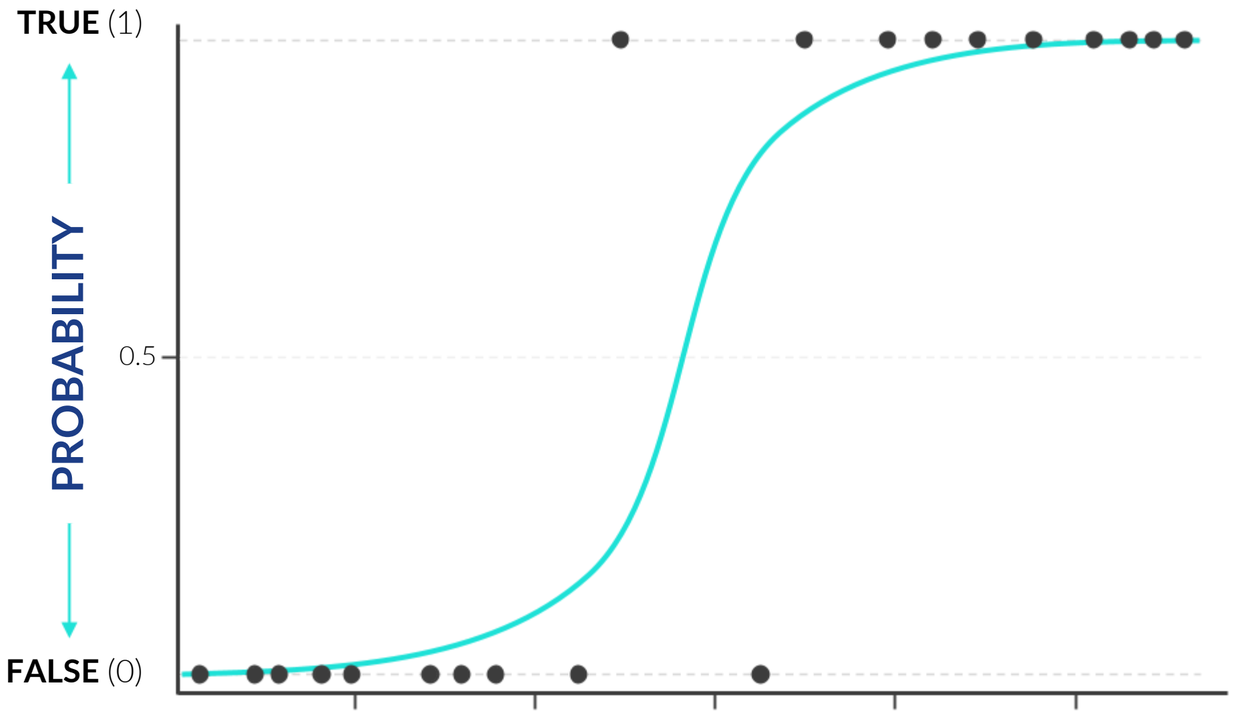
- Will this loan application be approved or denied?

LOGISTICS REGRESSION

Logistics Regression

Goals of Logistics

Evaluation

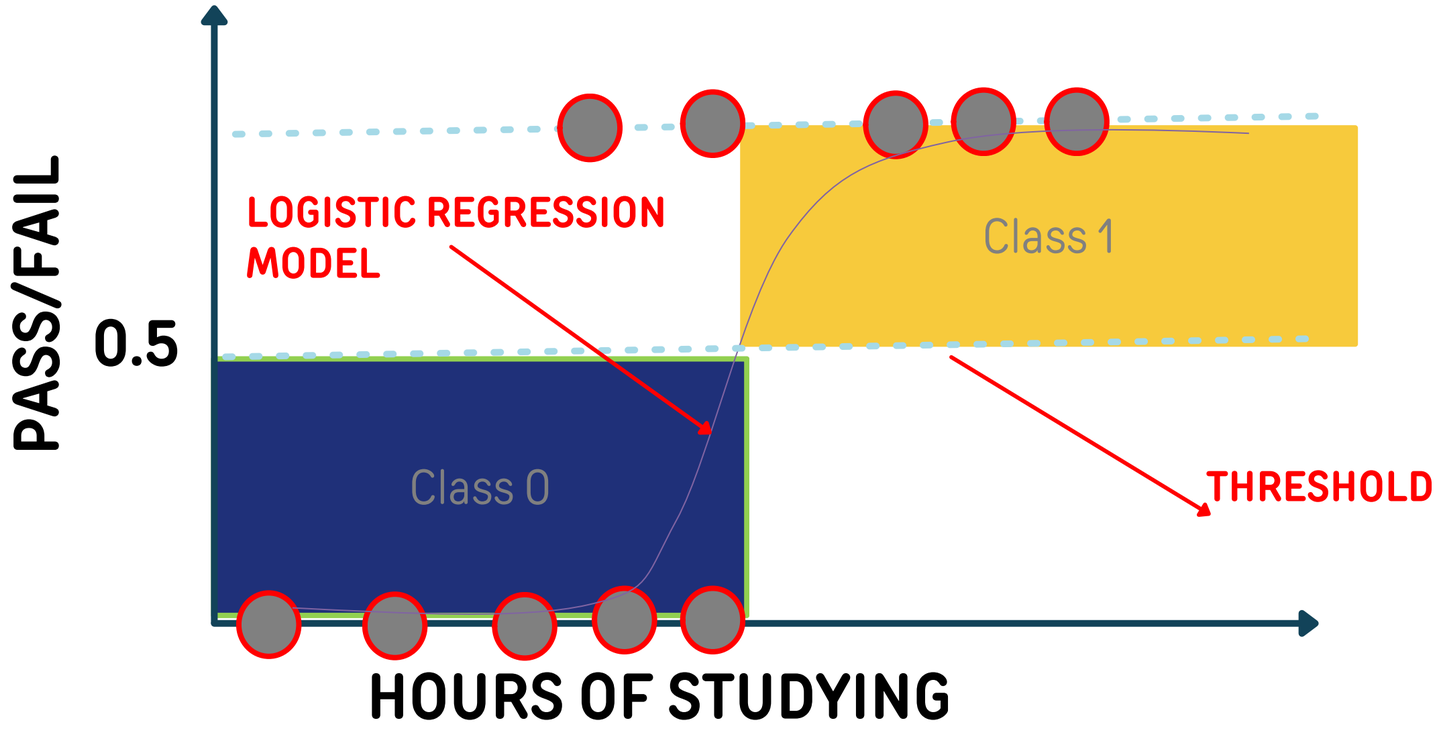


- Logistic regression plots the **best-fitting curve between 0 and 1**, which tells us the probability of Y being TRUE for any given value of X1

LOGISTICS REGRESSION

LOGISTIC REGRESSION: FROM PROBABILITY TO CLASS

- Now we need to convert from a probability to a class value which is “0” or “1”.



GOALS OF REGRESSION

ទិន្នាស្ថានសាន់រ៉ាយស៍
SUNRISE INSTITUTE

Regression models are used for two primary goals: **prediction** and **inference**

The goal shapes the modeling approach, including the regression algorithm used, the complexity of the model, and more

Logistics Regression

Goals of Logistics

Evaluation

PREDICTION

Used to **predict** the target as accurately as possible

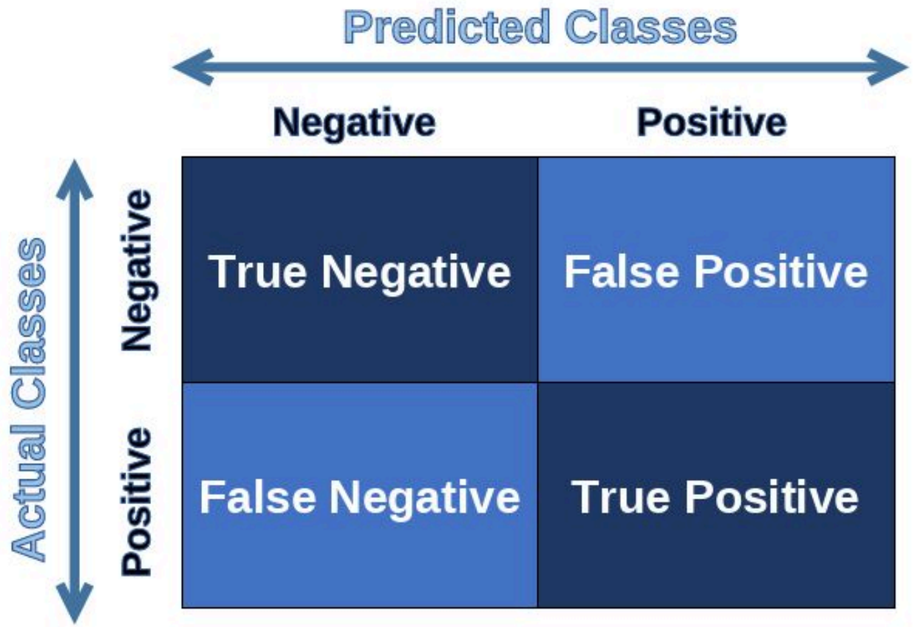
“What is the predict charges for a client given their age?”

INFERENCE

- Used to **understand the relationships** between the features and target
- *“How much do a age impact its charges?”*

What is Confusion Matrix?

A **confusion matrix** is defined as the table that is often used to describe the performance of a classification model on a set of the test data for which the true values are known.



PRECISION & RECALL

ទិន្នាស្ថានសាន់រ៉ាយស៍
SUNRISE INSTITUTE

Logistics Regression

Goals of Logistics

Evaluation

Precision indicates the percentage of positive predictions made by the model that are actually correct.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

Precision indicates the percentage of actual positive cases that the model correctly identifies

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

SPECIFICITY AND TNV

ទិន្នាស្ថានសាន់រ៉ាយស៍
SUNRISE INSTITUTE

Logistics Regression

Goals of Logistics

Evaluation

True Negative Values focuses on the proportion of correct negative predictions out of all predicted negatives.

$$\text{TNV} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Negatives}}$$

Specificity measures the ability of the model to identify all actual negative cases.

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$