# Confusion Matrix

### Sklearn Representation

Scikit learn documentation says — Wikipedia and other references may use a different convention for axes.

**A)**

| | | Actual Label 1 | Actual Label 0 |
|---|---|---|---|
| Predicted Label | 1 | TP | FP |
| | 0 | FN | TN |

**B)**

| | | Actual Label 0 | Actual Label 1 |
|---|---|---|---|
| Predicted Label | 0 | TN | FN |
| | 1 | FP | TP |

**C)**

| | | Predicted Label 1 | Predicted Label 0 |
|---|---|---|---|
| Actual Label | 1 | TP | FN |
| | 0 | FP | TN |

**D)**

| | | Predicted Label 0 | Predicted Label 1 |
|---|---|---|---|
| Actual Label | 0 | TN | FP |
| | 1 | FN | TP |

https://towardsdatascience.com/understanding-the-confusion-matrix-from-scikit-learn-c51d88929c79

# $F_\beta$-Score

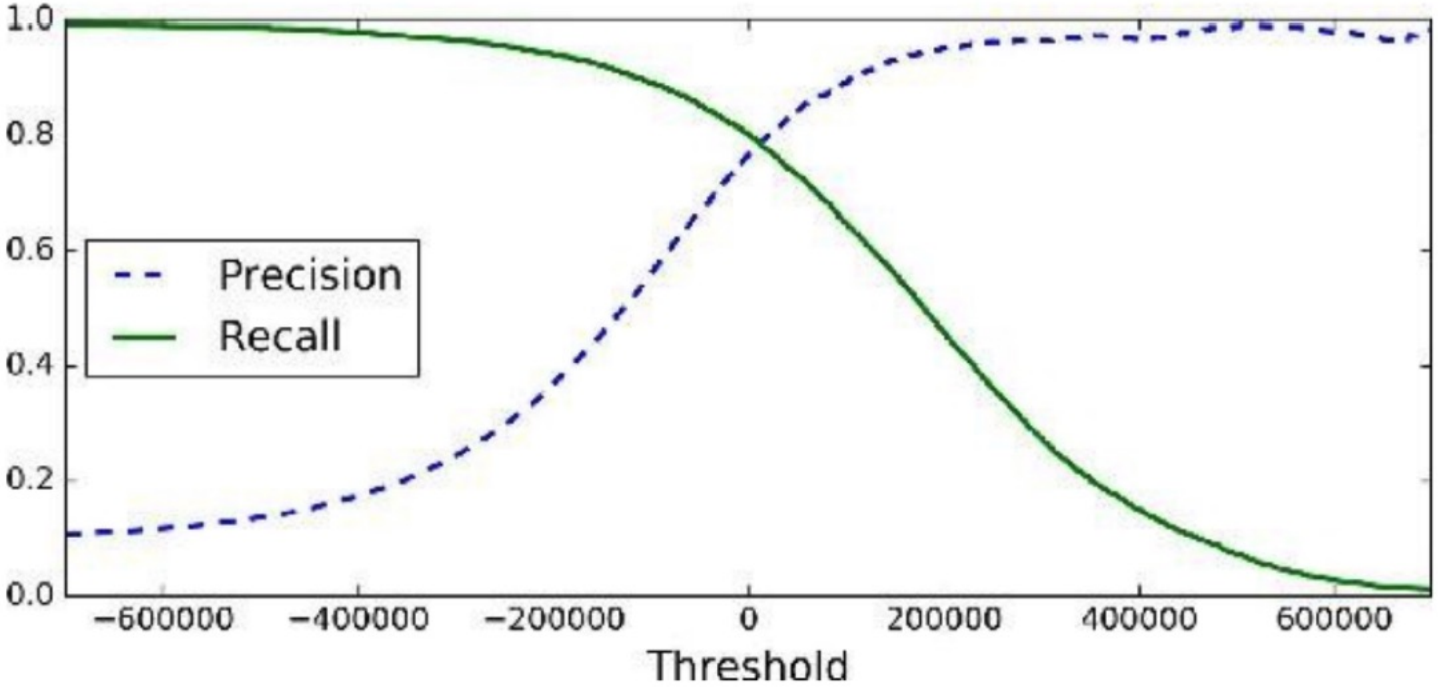
$$F_\beta = (1+\beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}}$$

$\beta = 1$

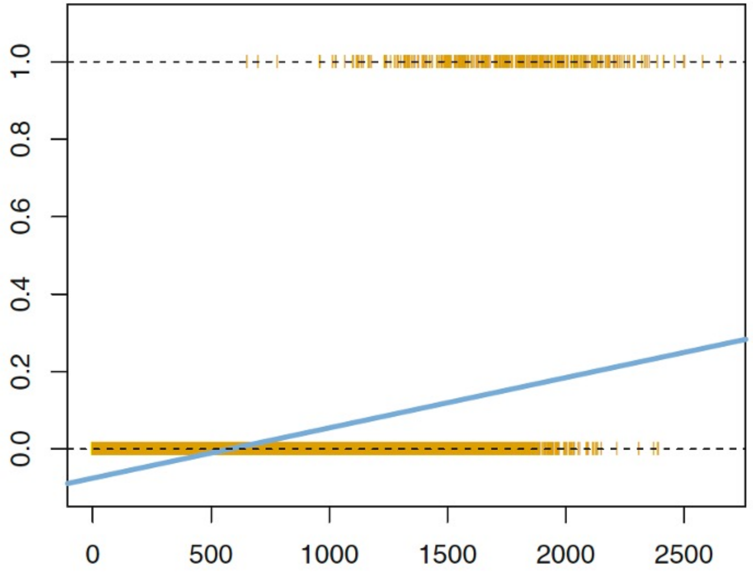
$$F_1 = \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}} = 2\frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} = \frac{\text{tp}}{\text{tp} + \frac{1}{2}(\text{fp} + \text{fn})}$$

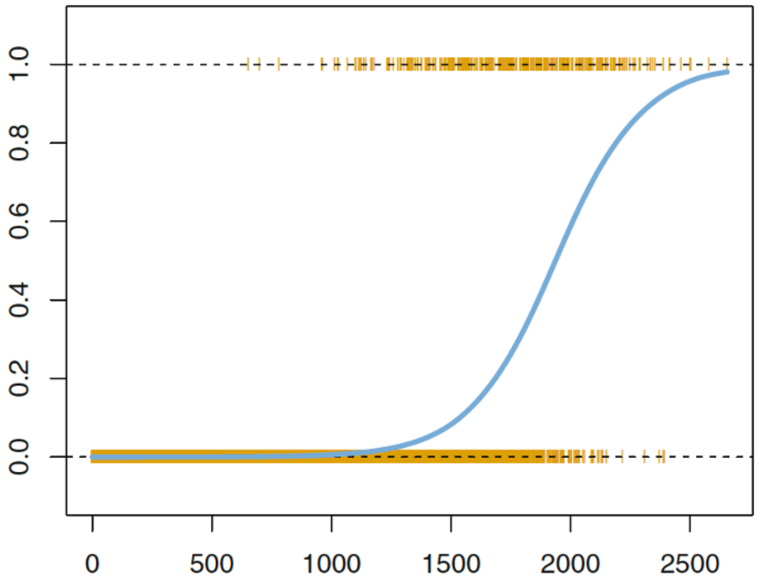
*harmonic mean*

# Precission – Racall Trade off



# Logistic Regression




$$p(X) = \beta_0 + \beta_1 X$$

$$p(X) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$$

# Logistic Regression

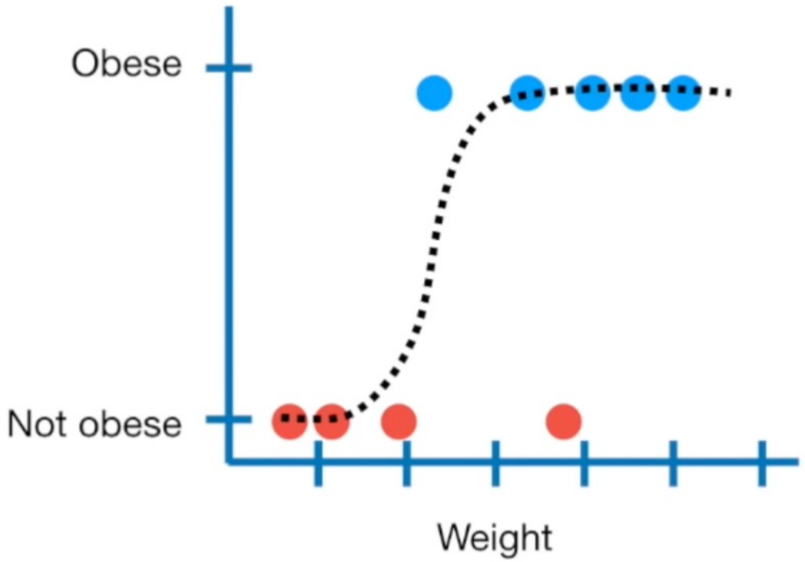
$$p(X) = \frac{e^{\beta_0+\beta_1 X}}{1 + e^{\beta_0+\beta_1 X}}$$

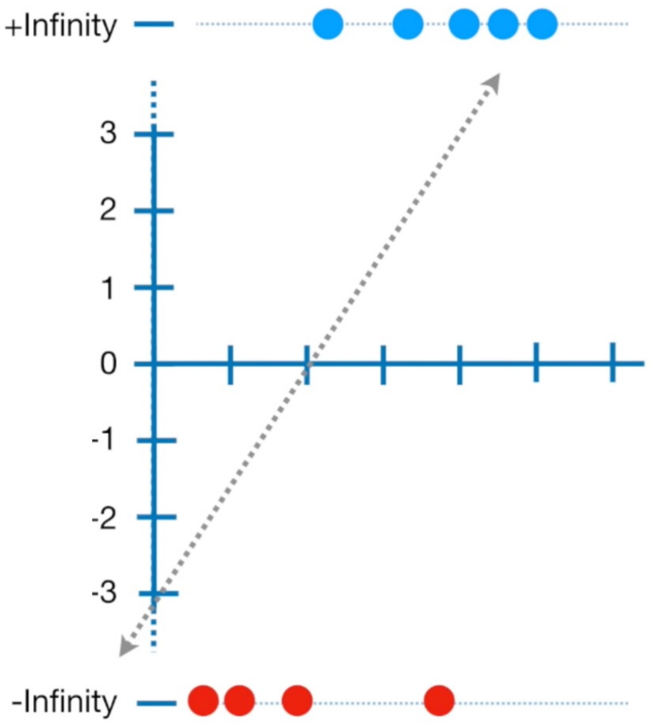
$$\frac{p(X)}{1 - p(X)} = e^{\beta_0+\beta_1 X}$$

$$\log\left(\frac{p(X)}{1 - p(X)}\right) = \beta_0 + \beta_1 X$$

*(log odds)*

# Logistic Regression

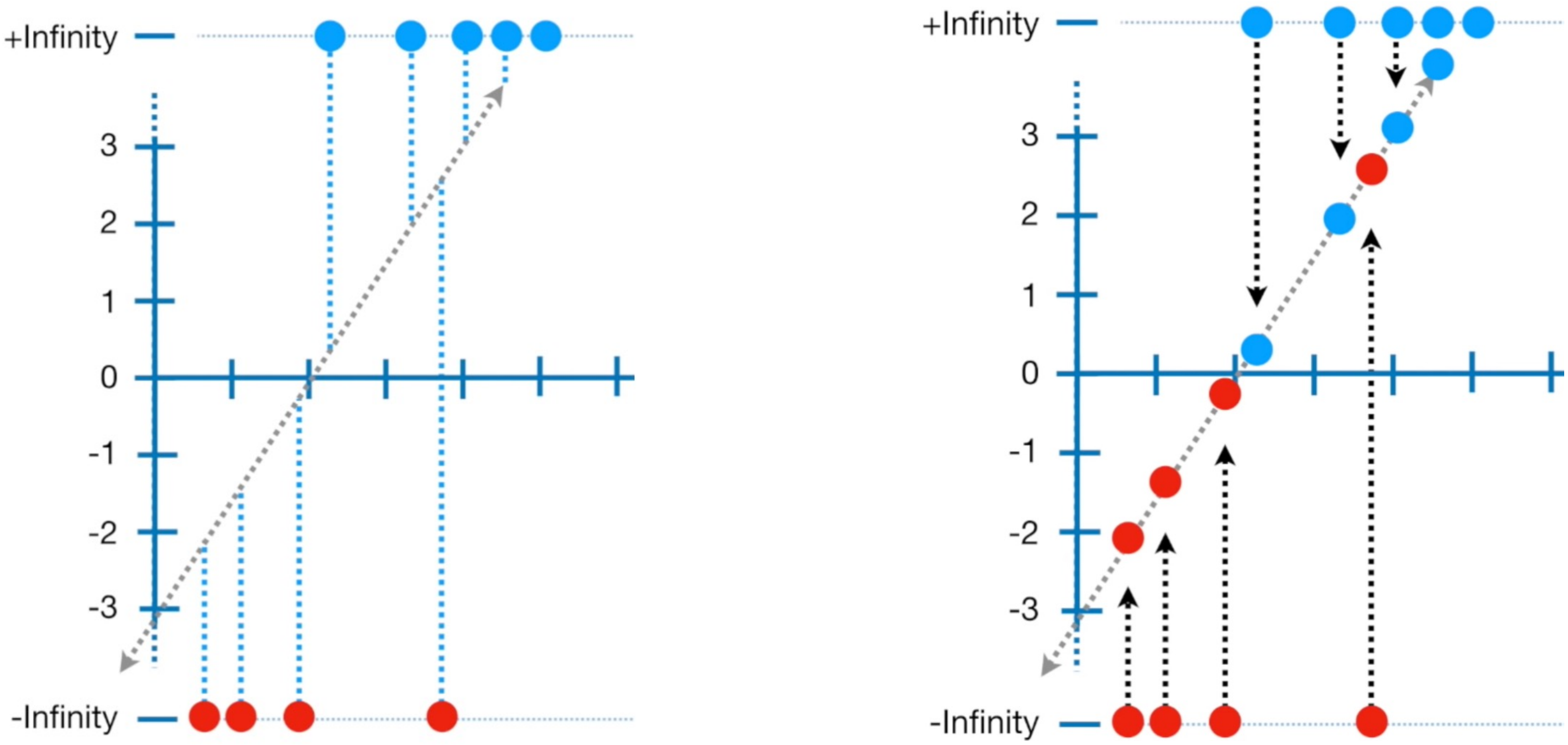


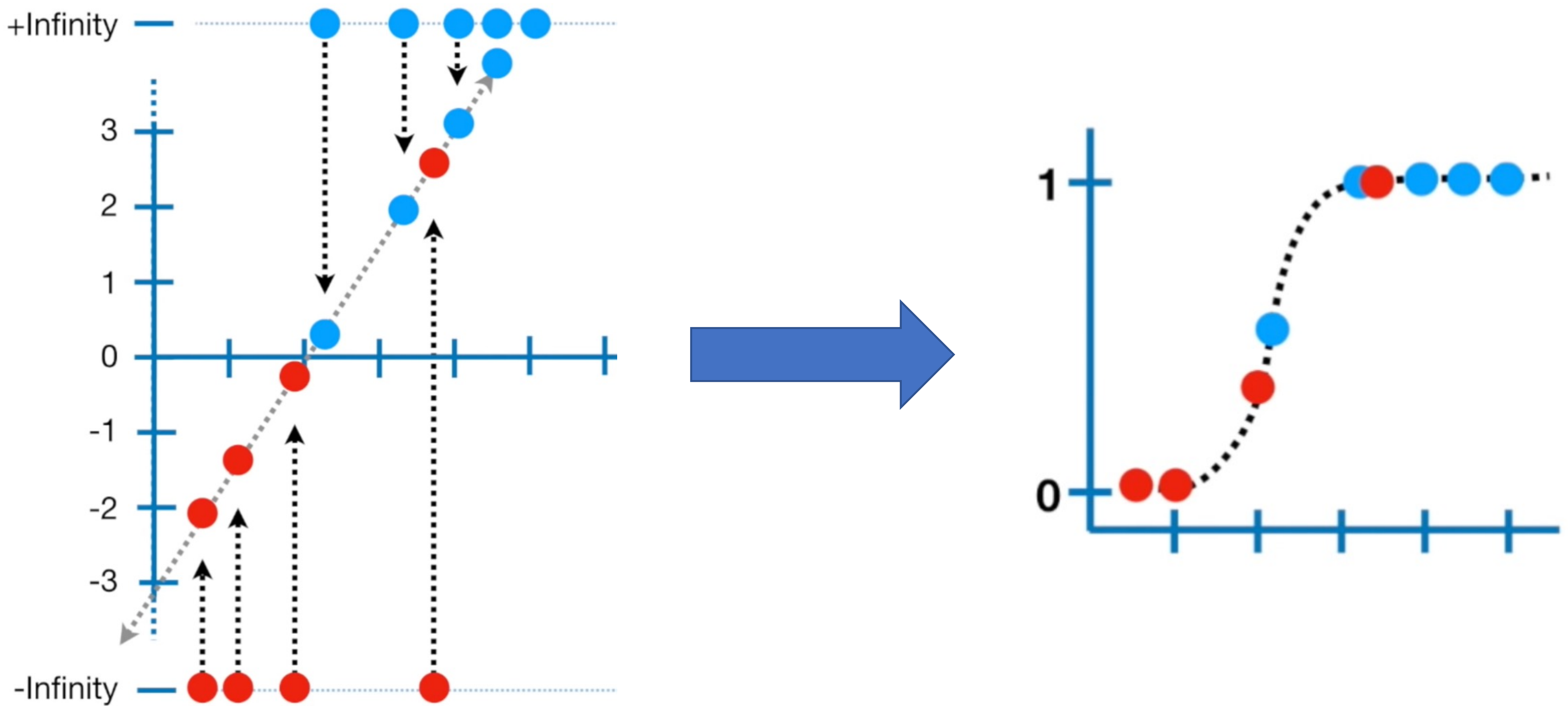

$$p(X) = \frac{e^{\beta_0+\beta_1 X}}{1 + e^{\beta_0+\beta_1 X}}$$

$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X$$

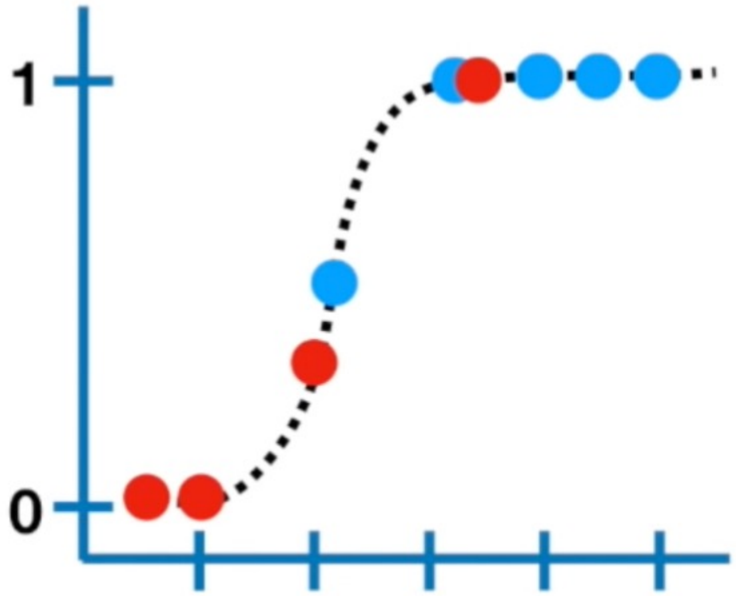
# Logistic Regression



# Logistic Regression



# Logistic Regression




Likelihood:

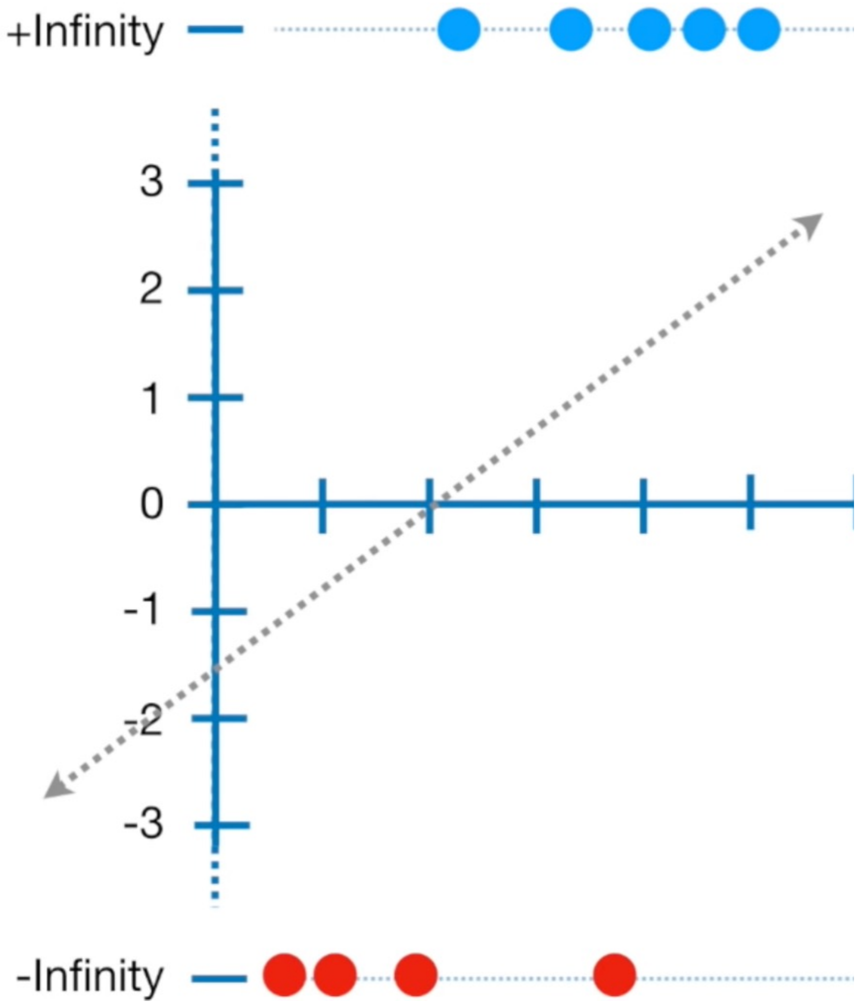
$$\ell(\beta_0, \beta_1) = \prod_{i:y_i=1} p(x_i) \prod_{i^i:y_{i'}=0} (1 - p(x_{i'}))$$

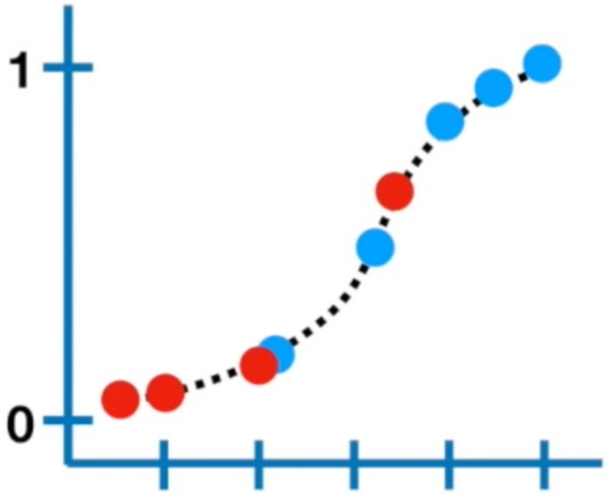
Log-likelihood:

$$\log \ell(\beta_0, \beta_1) = \sum_{i:y_i=1} \log p(x_i) + \sum_{i^i:y_{i'}=0} \log(1 - p(x_{i'}))$$

= log(0.49) + log(0.9) + log(0.91) + log(0.91) +
log(0.92) + log(1 - 0.9) + log(1 - 0.3) +
log(1 - 0.01) + log(1 - 0.01)

# Logistic Regression




$= \log(0.22) + \log(0.4) + \log(0.8) + \log(0.89) +$
$\log(0.92) + \log(1 - 0.6) + \log(1 - 0.2) +$
$\log(1 - 0.1) + \log(1 - 0.05)$

# Logistic Regression – Loss function

Linear Regression:

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

$$\mathrm{Cost}(h_\theta(x^{(i)}), y^{(i)}) = \frac{1}{2}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

non – convex

Logistic Regression:

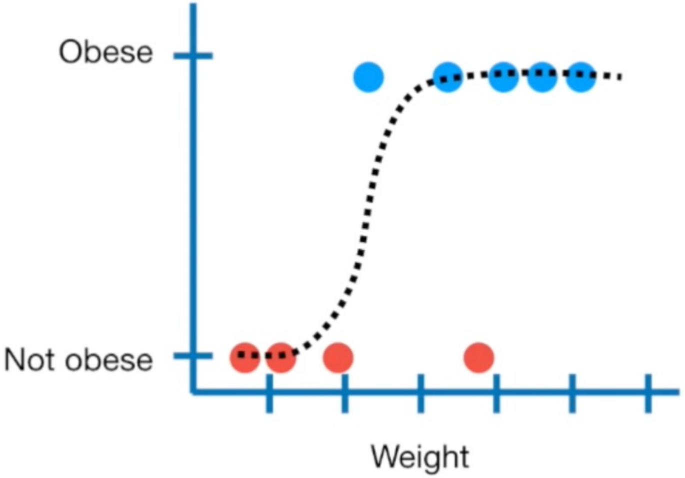
$$\mathrm{Cost}(h_\theta(x), y) = \begin{cases} -\log(h_\theta(x)) & \text{if } y = 1 \\ -\log(1 - h_\theta(x)) & \text{if } y = 0 \end{cases}$$

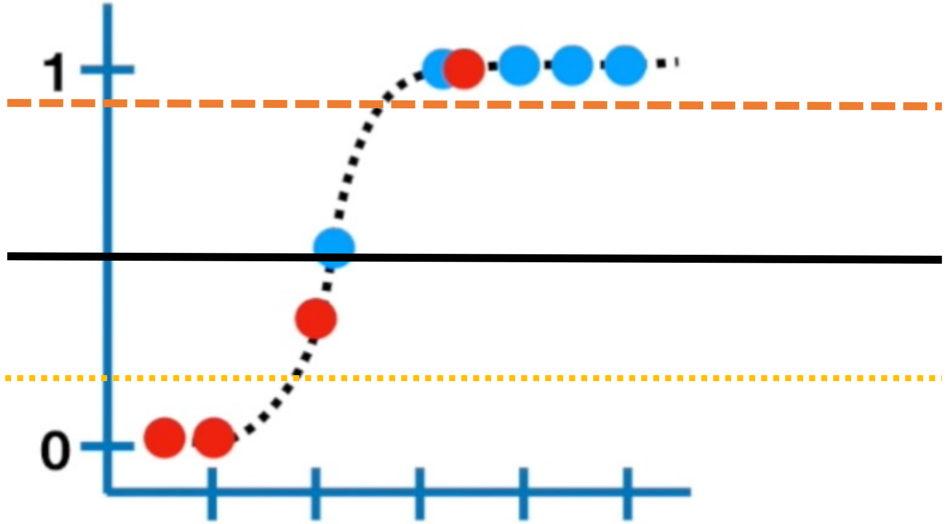
$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\mathrm{Cost}(h_\theta(x^{(i)}), y^{(i)})$$

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

cross-entropy

# ROC-AUC



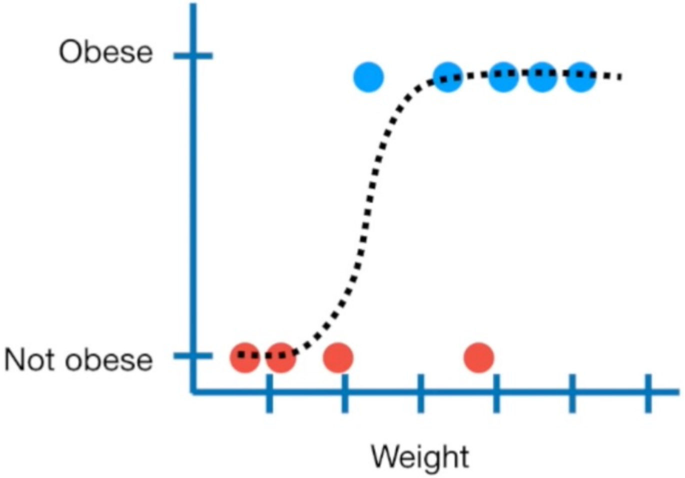



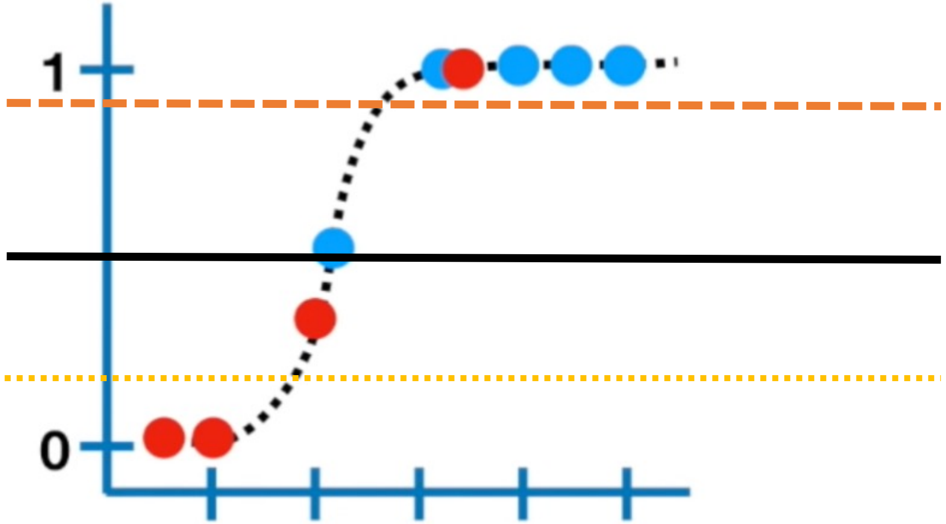

| | | ACTUALS 1 | ACTUALS 0 |
|---|---|---|---|
| PREDICTIONS | 1 | | |
| | 0 | | |

| | | ACTUALS 1 | ACTUALS 0 |
|---|---|---|---|
| PREDICTIONS | 1 | | |
| | 0 | | |

| | | ACTUALS 1 | ACTUALS 0 |
|---|---|---|---|
| PREDICTIONS | 1 | | |
| | 0 | | |

# ROC-AUC



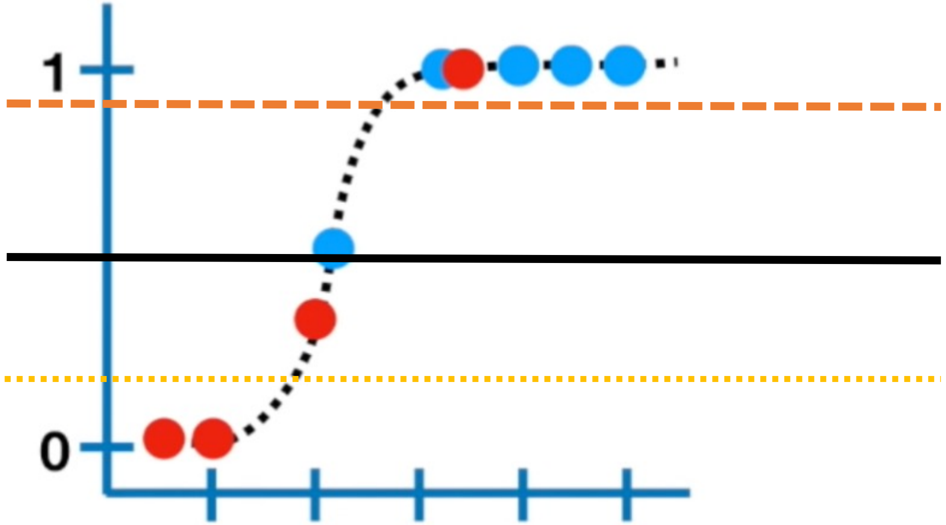




| | | ACTUALS | |
|---|---|---|---|
| | | 1 | 0 |
| PREDICTIONS | 1 | 4 | 1 |
| | 0 | 1 | 3 |

| | | ACTUALS | |
|---|---|---|---|
| | | 1 | 0 |
| PREDICTIONS | 1 | 5 | 0 |
| | 0 | 1 | 3 |

| | | ACTUALS | |
|---|---|---|---|
| | | 1 | 0 |
| PREDICTIONS | 1 | 5 | 0 |
| | 0 | 2 | 2 |

# ROC-AUC




| PREDICTIONS \ ACTUALS | 1 | 0 |
|---|---|---|
| 1 | 4 | 1 |
| 0 | 1 | 3 |

$$TPR = \frac{TP}{TP + FN} = \frac{4}{4 + 1}$$

$$FPR = 1 - TPR = \frac{FP}{FP + TN} = \frac{1}{1 + 3}$$

| PREDICTIONS \ ACTUALS | 1 | 0 |
|---|---|---|
| 1 | 5 | 0 |
| 0 | 1 | 3 |

$$TPR = \frac{5}{5 + 1}$$

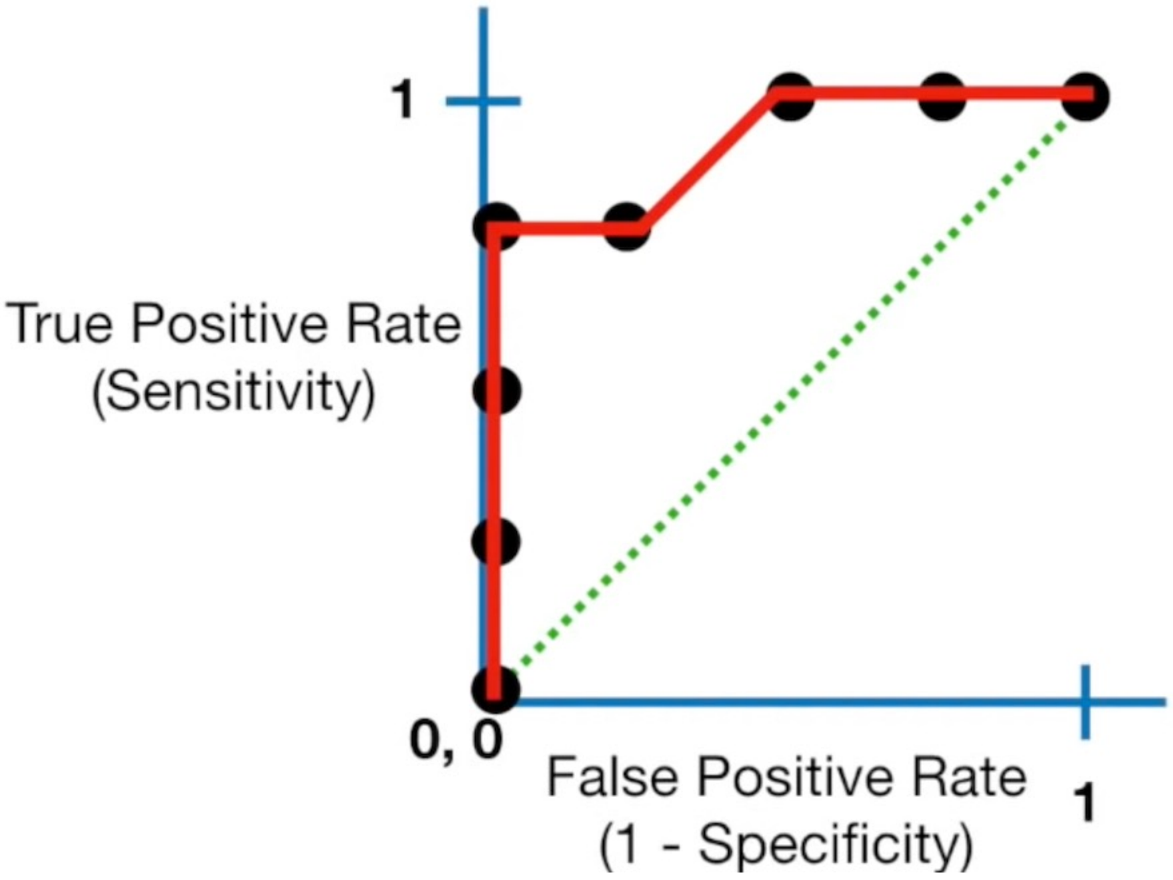
$$FPR = \frac{0}{0 + 3}$$

| PREDICTIONS \ ACTUALS | 1 | 0 |
|---|---|---|
| 1 | 5 | 0 |
| 0 | 2 | 2 |

$$TPR = \frac{5}{5 + 2}$$

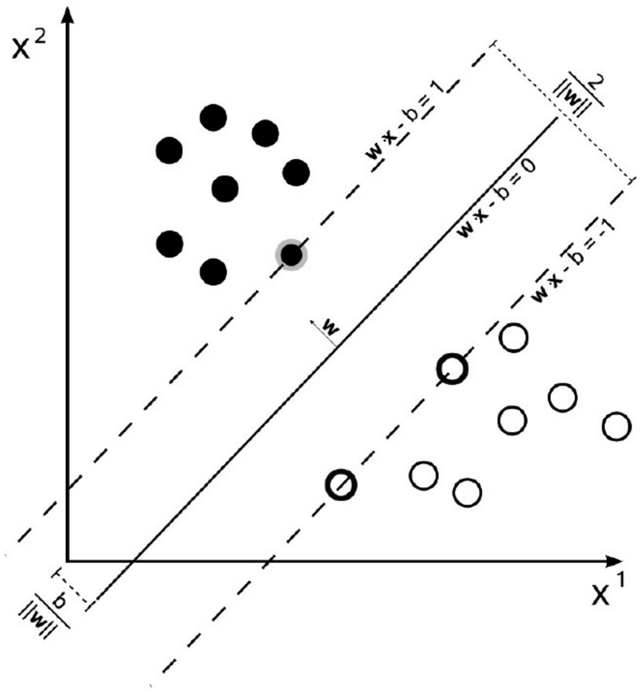
$$FPR = \frac{0}{0 + 2}$$

# ROC-AUC



1
True Positive Rate
(Sensitivity)
0, 0
False Positive Rate
(1 - Specificity)
1

# SVM – Hard Margin




$$f(x) = w^T x - b$$

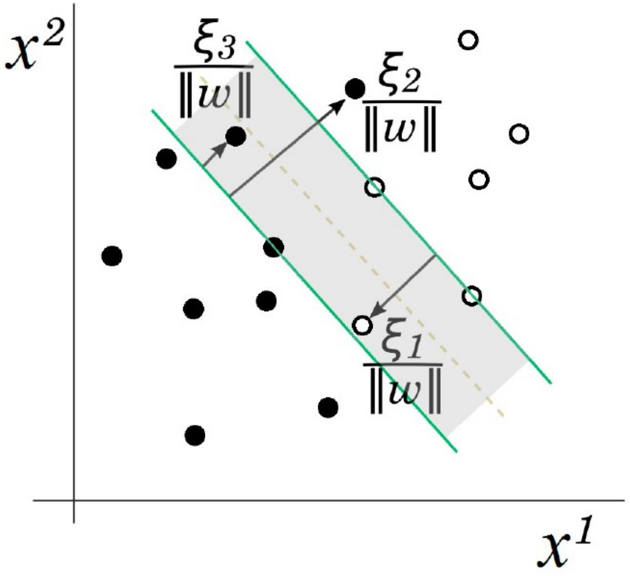
$$(w, b) = \underset{w, b}{\arg\min} \|w\|^2$$

Constraints:

$$w^T x_i - b \geq 1, \quad x_i \in Class$$
$$w^T x_i - b \leq -1, \quad x_i \notin Class$$

$$\Rightarrow \quad (w^T x_i - b) y_i \geq 1$$

# SVM – Soft Margin




$$\xi_i = \max\{1 - f(x_i)\, y_i, 0\}$$

Hinge loss

$$f(x) = w^T x - b$$

$$(w, b) = \underset{w,b}{\arg\min} \sum_i \xi_i + \lambda \|w\|^2$$

$$\lambda > 0$$

Constraints, for each $i$:

$$\xi_i \geq 0$$
$$(w^T x_i - b)\, y_i \geq 1 - \xi_i$$

# Decission Tree - CART

MACHINE LEARNING MASTERY

https://machinelearningmastery.com

towards data science

https://towardsdatascience.com

https://www.youtube.com/c/joshstarmer

https://www.youtube.com/c/3blue1brown

https://www.youtube.com/c/TechWorldwithNana

https://www.youtube.com/c/TechWithTim

kaggle

Machine Learning Study Groups

https://www.youtube.com/channel/UCMEQFEKrsRFBXnUIreTACxg

TensorFlow

https://www.youtube.com/c/TensorFlow