Probabilistic discriminative models:

$\underline{w} \in \mathbb{R}^D$

$$P(c_1 \mid \underline{x}) = \sigma(\underline{w}^T \underline{x})$$

Computes '$\underline{w}$' using ML approach for D parameters is called "Logistic regression"

Goal: Given $\{\underline{x}_i, y_i\}_{i=1}^{N}$ with $y_i \in \{0,1\}$

$$P(y \mid \underline{x}; \underline{w}) = \mathrm{Ber}(y \mid \sigma(\underline{w}^T x))$$

$$P(\underline{y} \mid \underline{x}; \underline{w}) = \prod_{i=1}^{N} P(y_i \mid \underline{x}_i; \underline{w})$$

$$= \prod_{n=1}^{N} [\sigma(\underline{w}^T \underline{x})]^{y_n} [1 - \sigma(\underline{w}^T \underline{x})]^{1-y_n}$$

<u>Negative likelihood function:</u>

$$f(\underline{w}) := - \sum_{n=1}^{N} \Big[y_n \log \sigma(\underline{w}^T \underline{x})$$

$$+ (1 - y_n) \log (1 - \sigma(\underline{w}^T \underline{x})) \Big]$$

Cross-entropy function

$$\sigma(-t) = 1 - \sigma(t)$$

Alternative view:

$$y_i \in \{0, 1\}$$

$$\tilde{y}_i = 2y_i - 1$$

$$\tilde{y}_i \in \{-1, +1\}$$

$$\boxed{\mathcal{L}(\underline{w}) = \sum_{i=1}^{N} \log\left(1 + \exp\left(-\tilde{y}_i \underline{w}^T x_n\right)\right)}$$

When

$$y_i = 1$$

$$-\log \sigma(\underline{w}^T \underline{x}) = \log\left(1 + \exp(-\underline{w}^T x_i)\right)$$

$$= \log\left(1 + \exp\left(-\tilde{y}_i \underline{w}^T x_i\right)\right)$$

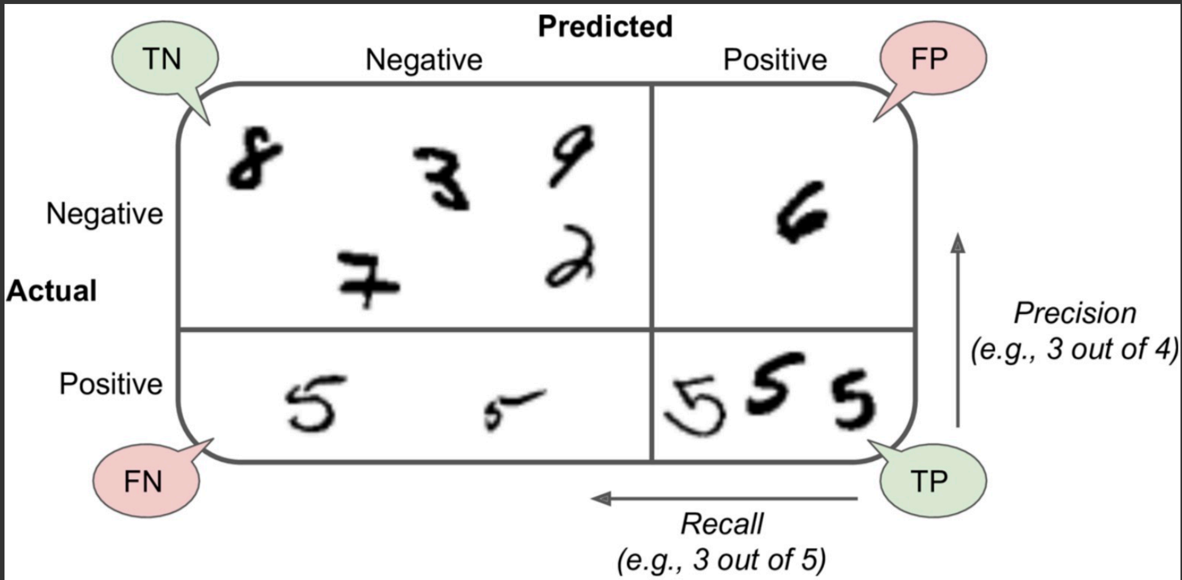
$$y_i = 0$$

$$-\log\left(1 - \sigma(\underline{w}^T \underline{x}_i)\right) = \log\left(1 + \exp(\underline{w}^T \underline{x}_i)\right)$$

$$= \log\left(1 + \exp\left(-\tilde{y}_i \underline{w}^T x_i\right)\right)$$

Multiclass logistic regression:

$$f(w) = -\sum_{n=1}^{N} \sum_{k=1}^{C} y_{nk} \log \frac{\exp(w_k^T x_n)}{\sum_j \exp(\underline{w}_j^T x_n)}$$



Accuracy:

$$\frac{TP + TN}{TP + TN + FP + FN}$$

Not very useful with class imbalance [990 +ve, 10 -ve]

99%

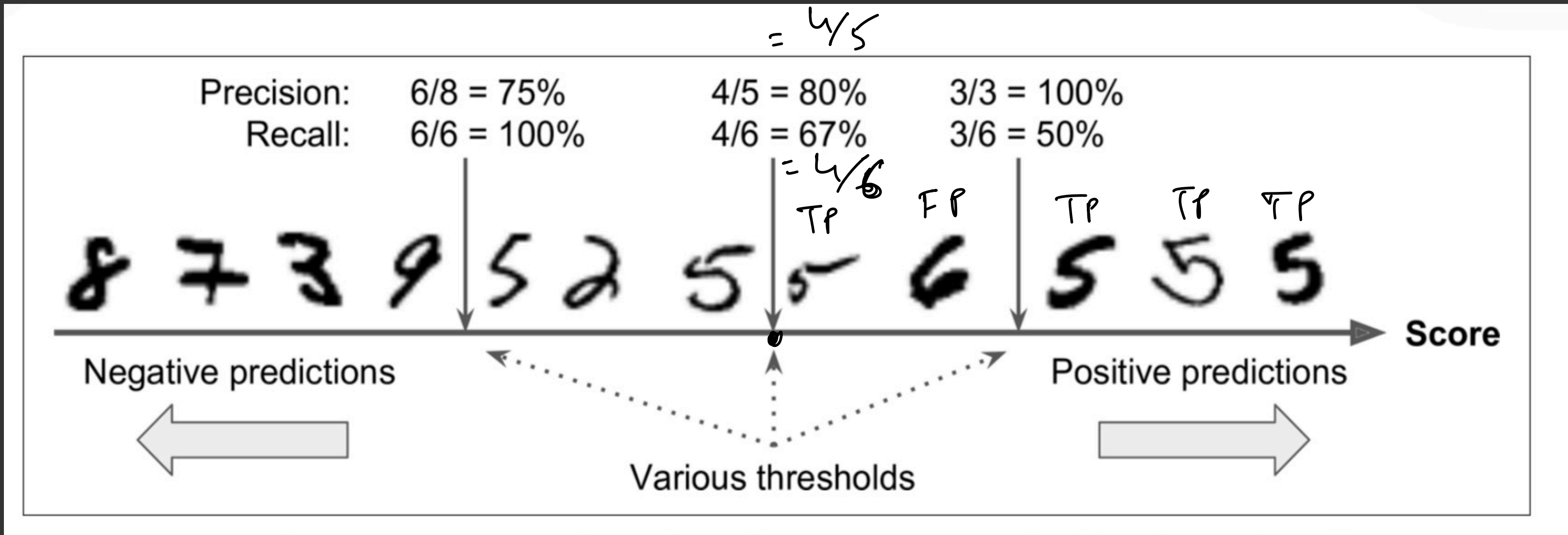
$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

(harmonic mean)

Trade off :


= 4/5
Precision: 6/8 = 75% 4/5 = 80% 3/3 = 100%
Recall: 6/6 = 100% 4/6 = 67% 3/6 = 50%
= 4/6
TP FP TP TP TP
Score
Negative predictions
Positive predictions
Various thresholds