- Boxplot: $Q_1 = N \times 0.25 \quad Q_2 = N \times 0.5 \quad Q_3 = N \times 0.75$
  $IQR = Q_3 - Q_1$ Bounds=$[Q_1 - 1.5 \times IQR, Q_3 + 1.5 \times IQR]$
- Pearson $= \dfrac{\Sigma(y_{1_i} - \bar{y}_1)(y_{2_i} - \bar{y}_2)}{\sqrt{\Sigma(y_{1_i} - \bar{y}_1)^2 \times (y_{2_i} - \bar{y}_2)^2}}$
- Spearman: Atribuir ranks e aplicar fórmula de Pearson.
  Exemplo: [20, 10, 20, 30, 20] $\rightarrow$ [3, 1, 3, 5, 3]
- Information Gain: $IG(y_{out}|y_i) = E(y_{out}) - E(y_{out}|y_i)$
- Entropy: $E(y) = -\sum P(x_i)\log(P(x_i))$
- Normalization:
  - MinMax: $\dfrac{x_i - min}{max - min}$
  - Standardization: $\dfrac{x_i - \mu}{\sigma}$
- Bipolarization
  - Range (equal width)
  - Frequency (equal depth)
- Confusion Matrix
- Metrics:
  - Accuracy $= \dfrac{TP + TN}{total}$
  - Recall $= \dfrac{TP}{TP + FN}$
  - Precision $= \dfrac{TP}{TP + FP}$
  - $\text{F}_1 = \dfrac{TP}{TP + \frac{1}{2}(FP + FN)}$
- Error:
  - MSE $= \sum(Z - \hat{Z})^2$
  - RMSE $=\sqrt{MSE}$
  - MAE $= \sum|Z - \hat{Z}|$
- Decision trees:
  - Escolher feature com maior IG.
  - Dividir dataset segundo essa feature, criar folhas se necessário.
  - Repetir até ser impossível continuar.

  Prune:
- Bayes Rule: $P(C|x) = \dfrac{P(C)P(x|C)}{P(x)}$
- K-Nearest Neighbors:
  - Distances: (for n variables)
    - Manhattan: $\sum_{i=1}^{n}|y_{i_1} - y_{i_2}|$
    - Euclidean: $\sqrt{\sum_{i=1}^{n}(y_{i_1} - y_{i_2})^2}$
    - Cosine:
    - Hamming: #Differences
  - Escolher K mais próximos
  - Classificar usando a média se a variável for numérica, ou a moda se for categórica
  - Caso seja weighted, dividir pelo peso
- Regressions:
  - Linear: $W = (X^TX)^{-1}X^TZ$
  - Ridge: $W = (X^TX + \lambda I)^{-1}X^TZ$
- Perceptron:
  $\hat{Z} = a(W^TX)$, $a \leftarrow$ activation function
  Se $Z \neq \hat{Z} \longrightarrow W' = W + \eta(Z - \hat{Z})X$
- Neural Networks (MLP):
  - Forward: $x^{[0]} \rightarrow z^{[1]} = w^{[1]}x^{[0]} + b^{[1]} \rightarrow x^{[1]} = a\left(z^{[1]}\right) \rightarrow \ldots$
    $\rightarrow z^{[i]} = w^{[i]}x^{[i-1]} + b^{[i]} \rightarrow x^{[i]} = a\left(z^{[i]}\right) \rightarrow E$
  - Backward:
    - $\delta^{[last]} = \dfrac{\delta E}{\delta x^{[last]}} \circ \dfrac{\delta x^{[last]}}{\delta z^{[last]}}$
    - $\delta^{[i]} = \left(\dfrac{\delta z^{[i+1]}}{\delta x^{[i]}}\right)^T \cdot \delta^{[i+1]} \circ \dfrac{\delta x^{[i]}}{\delta z^{[i]}}$
    - $w^{[i]\prime} = w^{[i]} - \eta\dfrac{\delta E}{\delta w^{[i]}}$
    - $\dfrac{\delta E}{\delta w^{[i]}} = \delta^{[i]} \cdot \left(\dfrac{\delta z^{[i]}}{\delta w^{[i]}}\right)^T$
    - $b^{[i]\prime} = b^{[i]} - \eta\dfrac{\delta E}{\delta b^{[i]}}$
    - $\dfrac{\delta E}{\delta b^{[i]}} = \delta^{[i]}$
    - $\dfrac{\delta z^{[i+1]}}{\delta x^{[i]}} = w^{[i+1]} \qquad \dfrac{\delta z^{[i]}}{\delta w^{[i]}} = x^{[i-1]} \qquad \dfrac{\delta z^{[i]}}{\delta b^{[i]}} = 1$

Name	Activation function	$\dfrac{\delta x^{[i]}}{\delta z^{[i]}}$
Sigmoid	$\sigma(x) = \dfrac{1}{1+e^{-x}}$	$\sigma(z^{[i]})(1 - \sigma(z^{[i]}))$
Hyperbolic tangent	$\tanh(x)$	$1 - \tanh\left(z^{[i]}\right)^2$

Name	Error function	$\dfrac{\delta E}{\delta x^{[i]}}$
Squared Error	$\dfrac{1}{2}\left(x^{[i]} - t\right)^2$	$x^{[i]} - t$
Cross-entropy	$-\sum_{i=1}^{n} t_i \log\left(x_i^{[i]}\right)$	$x^{[i]} - t$