1 Distanțe

1. **Distanța Manhattan:**
$$L_1((x_1, x_2, x_3 \ldots, x_n), (y_1, y_2, y_3 \ldots, y_n)) = \sum_{i=1}^{n} |x_i - y_i|$$
2. **Distanța Euler:**
$$L_2((x_1, x_2, x_3 \ldots, x_n), (y_1, y_2, y_3 \ldots, y_n)) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$
3. **Distanța Minkowski:**
$$L_p((x_1, x_2, x_3 \ldots, x_n), (y_1, y_2, y_3 \ldots, y_n)) = \sqrt[p]{\sum_{i=1}^{n} |x_i - y_i|^p}$$

2 Metrici

2.1 Clasificare

1. **Precizie:** (DOAR cazul binar) $P = \frac{TP}{TP+FP}$
2. **Recall:** (DOAR cazul binar) $R = \frac{TP}{TP+FN}$
3. **Specificitate:** (DOAR cazul binar) $S = \frac{TN}{TN+FP}$
4. **Acuratețe:** $A = \frac{T}{T+F}$
5. F_β**-score** (DOAR cazul binar) $F_\beta = (1 + \beta^2) \cdot \frac{P \cdot R}{(\beta^2 \cdot P) + R}$

2.2 Regresie

1. **Corelația Kendall-Tau:** $\tau_a = \frac{P-Q}{\frac{n(n-1)}{2}}$
$P = |\{(i,j) : 1 \leq i < j \leq n, (x_i - x_j)(y_i - y_j) > 0\}|$ (numărul de perechi concordante)
$Q = |\{(i,j) : 1 \leq i < j \leq n, (x_i - x_j)(y_i - y_j) < 0\}|$ (numărul de perechi disconcordante)

3 Bayes

1. $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$
2. $P(A|B) = \frac{P(A \cap B)}{P(B)}$
3. Teoremă: Clasificatorul Bayes h_{Bayes} este optim. (i.e. $\text{error}_{\text{true}}(h_{\text{Bayes}}) \leq \text{error}_{\text{true}}(h), \forall h$)

4 Funcții Kernel

Definiție: O funcție finit pozitiv semi-definită $k : X \times X \to \mathbb{R}$ pentru care există o funcție de scufundare $\phi : x \in \mathbb{R}^m \to \phi(x) \in F$ cu F - spațiu Hilbert astfel încât se verifică relația
$$k(x,z) = \langle \phi(x), \phi(z) \rangle$$

4.1 Exemple

1. **RBF (Gaussiană)**: $k(x,z) = e^{-\frac{||x-z||^2}{2\sigma^2}}$
2. **Intersecție**: $k(x,z) = \sum\limits_i \min\{x_i, z_y\}$
3. **Hellinger**: $k(x,z) = \sum\limits_i \sqrt{x_i \cdot z_i}$
4. **PQ**: $k(x,z) = 2(P-Q)$

4.2 Obținerea de noi funcții kernel

(i) $k(x,z) = k_1(x,z) + k_2(x,z)$

(ii) $k(x,z) = ak_1(x,z)$

(iii) $k(x,z) = k_1(x,z) \cdot k_2(x,z)$

(iv) $k(x,z) = f(x) \cdot f(z)$

(v) $k(x,z) = x'Bz$

$k1, k2$ - funcții kernel
a - constantă reală pozitivă
f - funcție cu valori reale
B - matrice simetrică pozitiv semidefinită

4.3 Normalizarea datelor

1. **Forma primală**: $\frac{\phi(x)}{||\phi(x)||}$
2. **Forma duală**: $\frac{k(x_i,z_j)}{\sqrt{k(x_i,z_i) \cdot k(x_j,z_j)}}$
3. **Matricea kernel**: $\frac{K_{ij}}{\sqrt{K_{ii} \cdot K_{jj}}}$
4. **Forma duală a unei matrici** X: $X \cdot X^T$

5 Regresii

1. **Costul regresiei Lasso**: $\text{cost}(y, yhat) = \sum\limits_i (yhat_i - y_i)^2 + \alpha||W||^1$
2. **Costul regresiei Ridge**: $\text{cost}(y, yhat) = \sum\limits_i (yhat_i - y_i)^2 + \alpha||W||^2$

6 Erori

1. **Mean Squared Error (MSE)**: $MSE(y, yhat) = \frac{1}{n}\sum\limits_{i=1}^{n} (yhat_i - y_i)^2$
2. **Mean Absoulte Error (MAE)**: $MAE(y, yhat) = \frac{1}{n}\sum\limits_{i=1}^{n} |yhat_i - y_i|$

7 Calcule convoluții

- Input: $H \times W \times D$
- Filtre: x filtre de $y \times y$
- Stride: z
- Padding: p

1. **Număr parametrii strat**: $n_{\text{parametrii}} = (y \cdot y \cdot D + 1) \cdot x$
2. **Dimensiune output**: $O \times O \times D$, unde $O = (W - y + 2 \cdot p)/z + 1$
   (+1 vine de la bias)

8 Perceptroni

1. **Output**: $Y = X \cdot W + b$
   X - input, W - weights, b - bias.
2. **Reactualizare weights**: $NW = W - l \cdot G$
   W - weights, G - gradienți, l - learning rate

9 Funcții de activare

1. **ReLU**: $f(x) = \max(0, x) \in [0, \infty]$
2. **PReLU**: $f(x) = \begin{cases} \text{x} & x \geq 0 \\ \alpha \cdot x & x < 0 \end{cases}$
3. **Leaky ReLU**: $f(x) = \max(0.1x, x)$ 0.01
4. **ELU**: $f(x) = \begin{cases} \text{x} & x \geq 0 \\ \alpha(\exp(x) - 1) & x < 0 \end{cases}$
5. **Sigmoid**: $\sigma(x) = \frac{1}{1+e^{-x}} \in [0, 1]$
6. **Tanh**: $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \in [-1, 1]$
7. **Softmax**: $\mathcal{L}_i = -\log(\frac{e^{s_{y_i}}}{\sum_j e^{s_j}})$
   s_{y_i} - scorul etichetei y_i

10 k-NN

1. **Outputul unei regresii prin k-NN**: $\frac{1}{K} \sum_{i=1}^{K} y_i$
   y_i - eticheta vecinului i
2. **Blestemul dimensionalității**: Avem nevoie de 5^n date pentru spațiul $\mathbb{R}^n$.