Kaavat

12. marraskuuta 2022

MNIST numeroiden tunnistus k:n lähimmän naapurin (kNN) menetelmällä - käytetyt kaavat

Koska työ sisältää runsaasti matemaattisia kaavoja, joita on vaikea järkevästi esittää .md formaatissa tai koodin kommenteissa, kaikki käytetyt matemaattiset kaavat on esitetty tässä dokumentissa. Kaikissa kaavoissa käytetään $\mathbb{R}^2$ avaruutta merkinnällä (x, y).

Euklidinen etäisyys sekä pisteen ja pistejoukon välinen etäisyys

Kahden pisteen välinen euklidinen etäisyys:

$$d(a, b) = ||a - b|| = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$$

Pistejoukon ja pisteen välinen etäisyys

$$d(a, \mathcal{B}) = \min_{b \in \mathcal{B}} ||a - b||$$

Hausdorffin etäisyydet

Hausdorffin etäisyyksien laskemiseksi tarvitaan seuraavia suunnattuja etäisyyksiä:

$$d_1(\mathcal{A}, \mathcal{B}) = \min_{a \in \mathcal{A}} d(a, \mathcal{B}) \tag{1}$$

$$d_2(\mathcal{A}, \mathcal{B}) = {}^{50}K^{th}_{a \in \mathcal{A}} d(a, \mathcal{B}) \tag{2}$$

$$d_3(\mathcal{A}, \mathcal{B}) = {}^{75}K^{th}_{a \in \mathcal{A}} d(a, \mathcal{B}) \tag{3}$$

$$d_4(\mathcal{A}, \mathcal{B}) = {}^{90}K^{th}_{a \in \mathcal{A}} d(a, \mathcal{B}) \tag{4}$$

$$d_5(\mathcal{A}, \mathcal{B}) = \max_{a \in \mathcal{A}} d(a, \mathcal{B}) \tag{5}$$

$$d_6(\mathcal{A}, \mathcal{B}) = \frac{1}{N_a} \sum_{a \in \mathcal{A}} d(a, \mathcal{B}) \tag{6}$$

$$d_{6b}(\mathcal{A}, \mathcal{B}) = \sum_{a \in \mathcal{A}} d(a, \mathcal{B})$$

sekä seuraavia suuntaamattomia etäisyyksiä:

$$f_1(d(\mathcal{A},\mathcal{B}), d(\mathcal{B},\mathcal{A})) = \min(d(\mathcal{A},\mathcal{B}), d(\mathcal{B},\mathcal{A})) \tag{7}$$

$$f_2(d(\mathcal{A},\mathcal{B}), d(\mathcal{B},\mathcal{A})) = \max(d(\mathcal{A},\mathcal{B}), d(\mathcal{B},\mathcal{A})) \tag{8}$$

$$f_3(d(\mathcal{A},\mathcal{B}), d(\mathcal{B},\mathcal{A})) = \frac{d(\mathcal{A},\mathcal{B}) + d(\mathcal{B},\mathcal{A})}{2} \tag{9}$$

$$f_4(d(\mathcal{A},\mathcal{B}), d(\mathcal{B},\mathcal{A})) = \frac{N_a d(\mathcal{A},\mathcal{B}) + N_b d(\mathcal{B},\mathcal{A})}{N_a + N_b} \tag{10}$$

Näistä suuntaamattomista ja suunnatuista etäisyyksistä saadaan Hausdorffin etäisyydet:

directed distance	function			
	f_1	f_2	f_3	f_4
d_1	D_1	D_2	D_3	D_4
d_2	D_5	D_6	D_7	D_8
d_3	D_9	D_{10}	D_{11}	D_{12}
d_4	D_{13}	D_{14}	D_{15}	D_{16}
d_5	D_{17}	D_{18}	D_{19}	D_{20}
d_6	D_{21}	D_{22}	D_{23}	D_{24}