

# Distance between data points

Gruppe E, AI OLA 2

## Manhattan

Manhattan geometri er geometri hvor distancen mellem 2 punkter er defineret til at være den absolutte forskel af de 2 koordinater. Navnet kommer af hvordan taxachauffører skal navigere imellem de store boligblokke på Manhattan

Den matematiske formel er givet ved:

$$d_{\text{Manhattan}}(A, B) = |x_1 - y_1| + |x_2 - y_2| + \dots + |x_n - y_n| = \sum_{i=1}^n |x_i - y_i|$$

Så ved et eksempel med et koordinatsystem med punkterne A = (3, 5) & B = (1,1), ville Manhattan distancen måles med ligningen:

$$d_{\text{Manhattan}} = |3 - 1| + |5 - 1| = 2 + 4 = 6$$

## Euclidean

Euclidean geometri er geometri hvor distancen mellem 2 punkter måles ved den direkte distance mellem dem, altså i en lige linje. Den er også kendt som Pythagorean distance da den bruger pythagoras læresætning

Den matematiske formel er givet ved:

$$d_{\text{Euclidean}}(A, B) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Så ved et eksempel med et koordinatsystem med punkterne A = (3, 5) & B = (1,1), ville Euclidean distancen måles med ligningen:

$$d_{\text{Euclidean}} = \sqrt{(3 - 1)^2 + (5 - 1)^2} = \sqrt{4 + 16} = \sqrt{20} \approx 4.47$$

# Hamming

Hamming distancen måler forskellen med 2 data. Disse data kan være strings, integers, eller vektorer. Denne distance måles ved at tælle antallet af positioner hvor dataen er forskellig. Distancen kan eksempelvis bruges til at vise hvor mange ændringer der er nødvendige for at konvertere et stykke data om til et andet.

Den matematiske formel er givet ved:

$$d_{\text{Hamming}}(A, B) = \sum_{i=1}^n I(x_i \neq y_i)$$

Et eksempel kunne være med disse 2 givne binære strenge:

- 1011001
- 0001101

Her er der forskel ved position 1, 3 og 5. og derved bliver hamming distancen = 3.