

Şule Meşe

191502001

1.Mesafe Nedir ?

Mesafe veriler arasındaki benzerlik veya uzaklığın ölçütüdür.

1.1 Öklid mesafesi (Euclidean distance) : Matematikte pisagor bağlantısı kullanılarak bulunan iki nokta arasındaki **mesafe** ölçüm birimidir. Buna göre iki boyutlu düzlemde iki nokta arasındaki **mesafe** basitçe iki noktanın x ve y koordinatlarının ayrı ayrı farklarının hipotenüs'üne eşittir.

$\{x_1, x_2, \dots, x_n\}$
 $\{y_1, y_2, \dots, y_n\}$  Herhangi iki nokta

Öklid Mesafesi  $d = \sqrt{\sum_{i=1}^k (x_i - y_i)^2}$

(Şekil 1- Öklid Mesafesi Formülü)

1.2 Manhattan mesafesi : Manhattan mesafesi, iki vektörün mutlak olarak farklarının toplamıdır. Yani iki nokta (X1, Y1) ve (X2, Y2) olarak verilirse Manhattan mesafesi $|X1 - X2| + |Y1 - Y2|$ olarak hesaplanır.

$\{x_1, x_2, \dots, x_n\}$
 $\{y_1, y_2, \dots, y_n\}$  Herhangi iki nokta

Manhattan Mesafesi  $d = \sum_{i=1}^k |x_i - y_i|$

(Şekil 2- Manhattan Mesafesi Formülü)

1.3 Minkowski mesafesi : Minkowski mesafesi Öklid mesafesi ve Manhattan mesafesi'nin bir genelleştirilmesi ile oluşturulur. Burada eklenen p parametresine göre işlemler yapılır p=1 olduğunda, Minkowski mesafesi, Manhattan mesafesine eşit olur. p=2 olduğunda, Minkowski mesafesi, Öklid mesafesine eşit olur.

$$\begin{array}{l} \{x_1, x_2, \dots, x_n\} \\ \{y_1, y_2, \dots, y_n\} \end{array} \xrightarrow{\text{Herhangi iki nokta}}$$

$$\text{Minkowski Mesafesi} \xrightarrow{\text{}} d = \left(\sum_{i=1}^k (|x_i - y_i|)^p \right)^{\frac{1}{p}}$$

p=1 Manhattan
p=2 Öklid

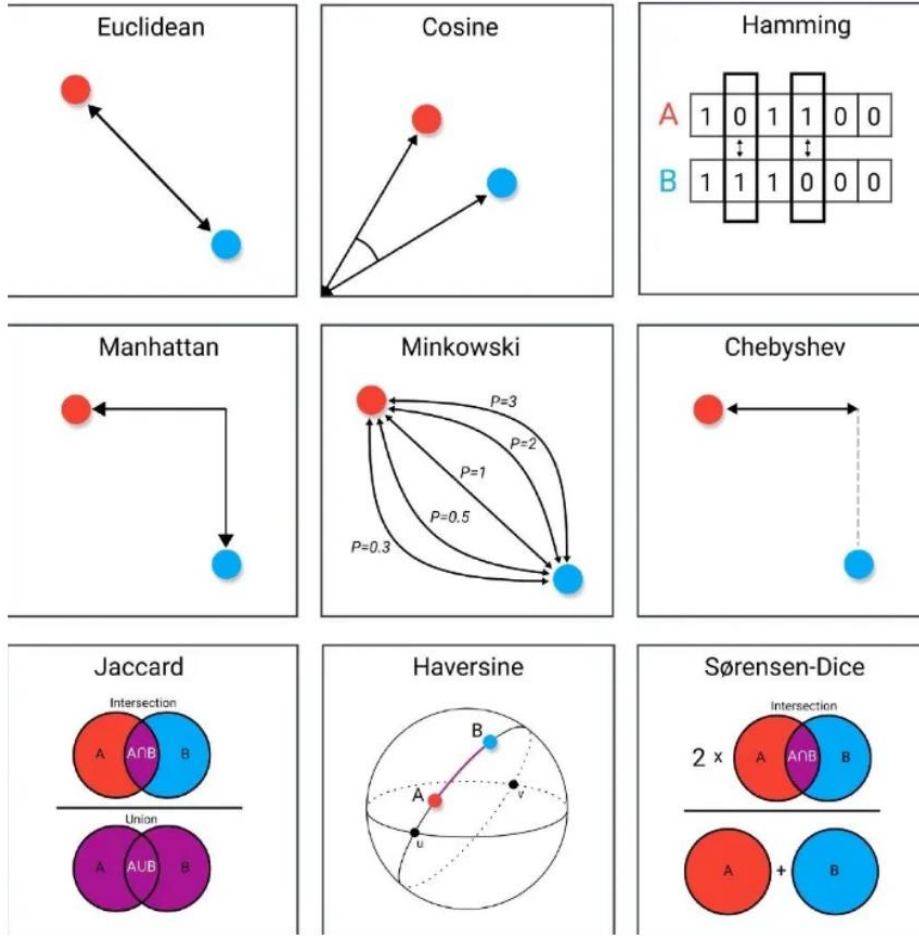
(Şekil 3 - Minkowski Mesafesi Formülü)

1.4 Kosinüs Benzerliği : Kosinüs benzerliği iki vektör arasındaki yön benzerliğini hesaplar. İki vektör arasındaki cosinüs açısı hesaplanır ve vektörlerin buna göre yönünün benzerliği belirlenir. Metin analizinde genellikle belge benzerliğini ölçmek için kullanılır. İki vektör aynı yöndeysse sonuç 1, ters yöndeysse -1 çıkar.

$$\| \mathbf{A} \| \| \mathbf{B} \| \quad \sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}$$

(Şekil 4- Kosinüs Benzerliği Mesafesi Formülü)

Hangi mesafe ölçüsünün kullanılacağı, veri türüne, uygulamaya ve probleme bağlı olarak değişebilir. Veri madenciliği ve makine öğrenimi projelerinde doğru mesafe ölçüsünü seçmek, sonuçların doğruluğunu ve performansını etkileyebilir.



Kaynakça

<https://www.cihanyakar.com/csharp-ile-vektor-normlari-ve-uzaklik-hesaplama/>

<https://bilgisayarkavramlari.com/2008/01/12/mesafe-vektoru-distance-vector/>

<https://bilisim.io/2023/07/03/vektor-veritabanlari-vector-databases/>

<https://abdullahatcili.medium.com/>