

1906003132015

# Doğal Dil İşleme

BAİBÜ Bilgisayar Müh.

Dr. Öğr. Üyesi İsmail Hakkı Parlak

[ismail.parlak@ibu.edu.tr](mailto:ismail.parlak@ibu.edu.tr)

Oda: 335

# Sayma (Counting) Vektörleri

- $d_1: [x, y, x, x] \rightarrow x:3, y:1$
- $d_2: [y, y, x] \rightarrow x:1, y:2$
- $d_3: [y, z, z, y, z, k] \rightarrow y:2, z:3, k:1$
  
- vocab: [unk, x, y, z, k]
- token indexleri: {unk: 0, x: 1, y: 2, z: 3, k: 4}
  
- $v_{d1}: [0, 3, 1, 0, 0]$
- $v_{d2}: [0, 1, 2, 0, 0]$
- $v_{d3}: [0, 0, 2, 3, 1]$

# Sayma (Counting) Vektörleri

- $d_4$ : [x, t, z, t, y, y]  $\rightarrow$  x:1, t:2, z:1, y:2
- vocab: [unk, x, y, z, k]
- token indexleri: {unk: 0, x: 1, y: 2, z: 3, k: 4}
- $v_{d4}$ : [2, 1, 2, 1, 0]

# Sayma (Counting) Vektörleri

- Veri seti matrisi  $(4 \times 5)$

- $v_{d1}$ :  $[0, 3, 1, 0, 0]$

- $v_{d2}$ :  $[0, 1, 2, 0, 0]$

- $v_{d3}$ :  $[0, 0, 2, 3, 1]$

- $v_{d4}$ :  $[2, 1, 2, 1, 0]$