Makine Öğrenmesi

RNN, attention ve pekiştirmeli öğrenme

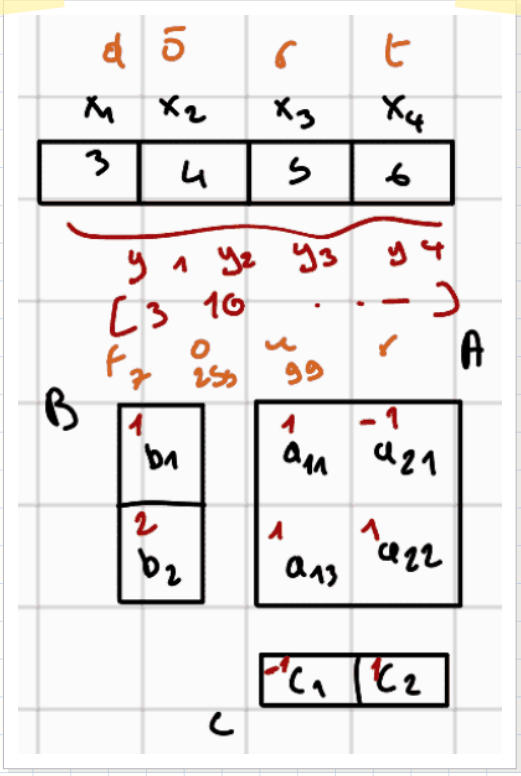
Orijinal not: Ders11.pdf

UYARI Bu belge, el yazısı ders notlarından Claude (Anthropic) yardımıyla dijitale aktarılmıştır. Yardımcı kaynak olarak hazırlanmıştır; hata veya eksik içerebilir. Lütfen orijinal notları ve ders kaynaklarını esas alın.

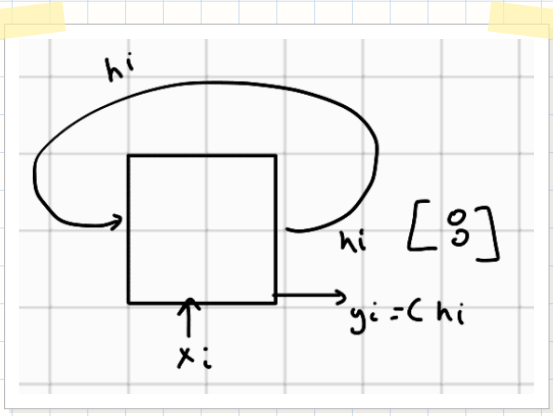
NOTICE *This document was transcribed from handwritten lecture notes with the help of Claude (Anthropic). It is meant as a supplementary resource and may contain errors or omissions. Always check against the original notes and official course material.*

orj. s. 1

RNN



orijinal çizim, s. 1



orijinal çizim, s. 1

$$\textcircled{1}\ [B \mid A]\begin{bmatrix} x_i \\ h_i \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = h_i = h_1$$

$Ch_i =$? okunamadı $= y_i$

$$\textcircled{2}\ \begin{bmatrix} 1 & 1 & -1 \\ 2 & 1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 17 \end{bmatrix}$$

$y_2 = 16$

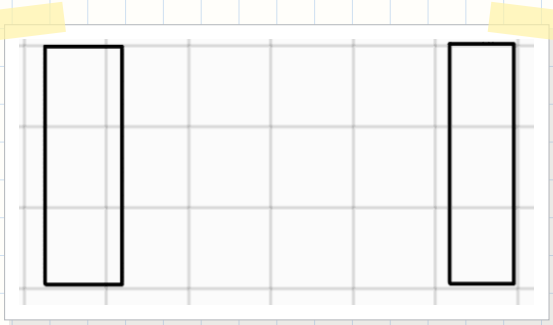
MSE =?

$$\hat{y} = f\left(CA\begin{bmatrix} x \\ h \end{bmatrix}\right)$$

$$L = (\hat{y} - y)^2$$

I run

I university bus ran



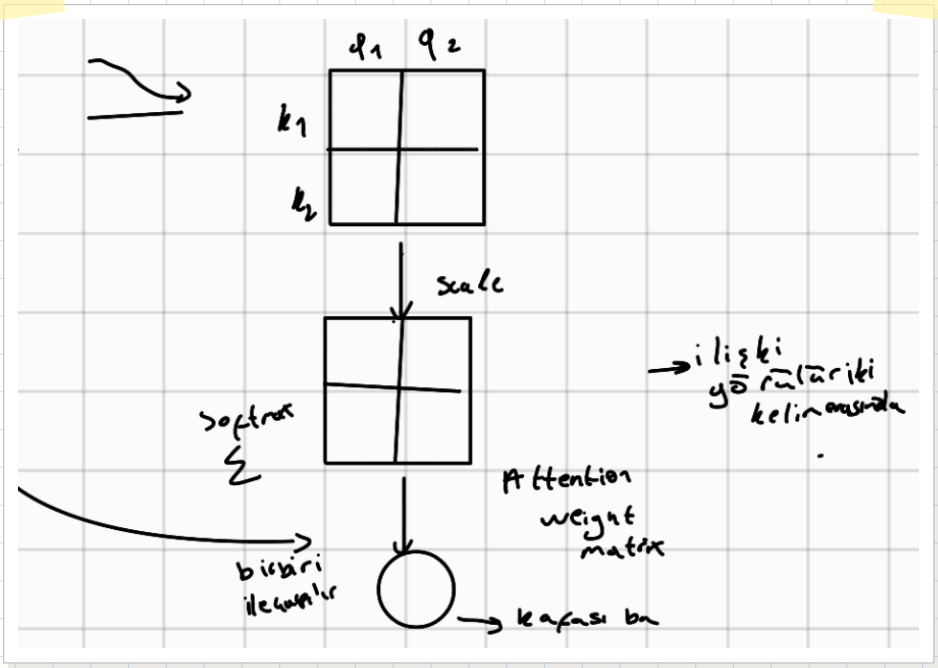
orijinal çizim, s. 1

Q = Query
k = key
V = Value

$$\text{attention} = \text{softmax}\left(\frac{k^T \cdot Q}{\underbrace{\sqrt{d_k}}_{k\text{'nın boyutu}}}\right) V$$

orj. s. 2

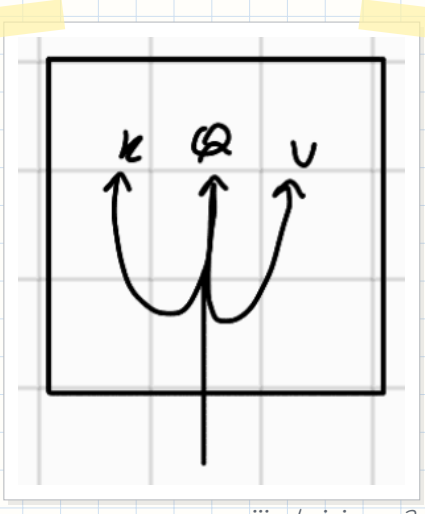
$W_Q \cdot w = q_1 q_2$
$W_K \cdot w = k_1, k_2$
$W_V \cdot w = v_1 v_2$



orijinal çizim, s. 2

Bunun adı self attention

Makalelerde böyle görürsünüz

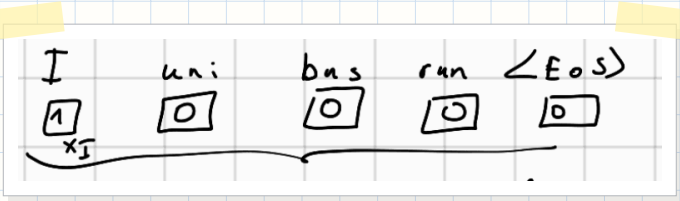


orijinal çizim, s. 2

→ multihead olabilir

Q : Ne aradığımız
k : Ne ile ilgili
V : Neyi kullanacağız

?sembol ile takım adını ?karşılaştırdık



orijinal çizim, s. 2

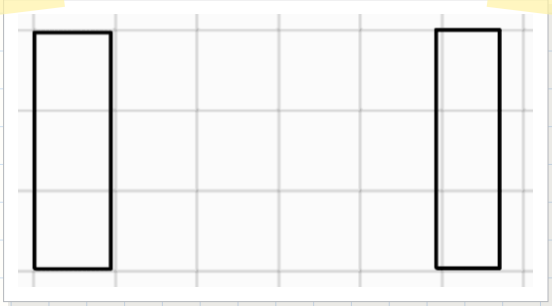
bunun $\times W_1 \rightarrow$ positional encoding matris

$$x_I + W_1 x_I = \underbrace{\bar{x}_I}_{\text{positional embedding}}$$

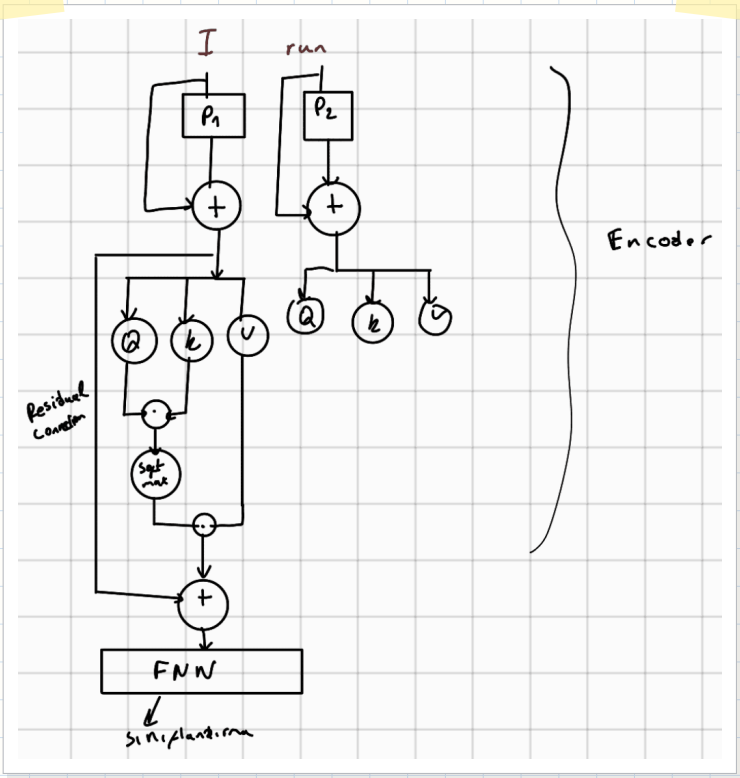
orj. s. 3

I run

I university bus run

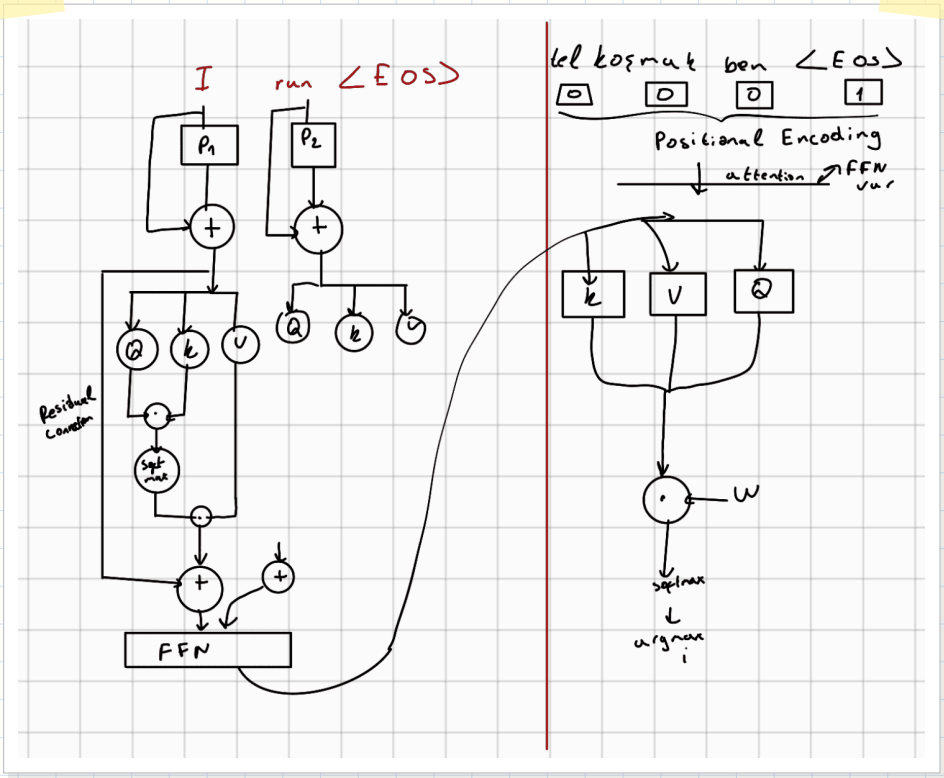


orijinal çizim, s. 3

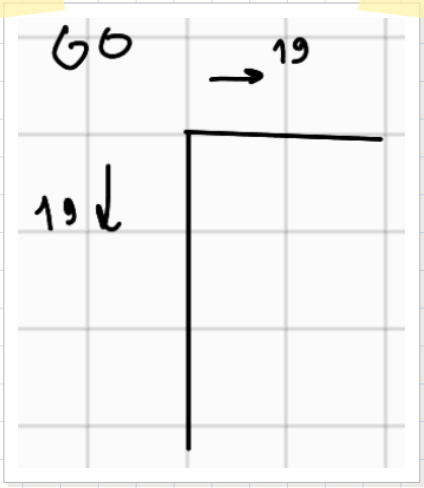


orijinal çizim, s. 3

orj. s. 4



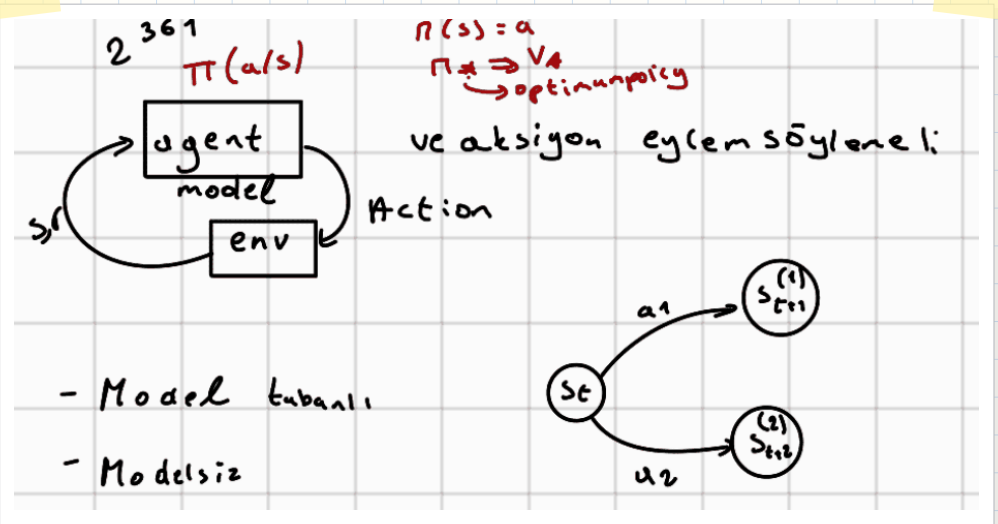
orijinal çizim, s. 4



orijinal çizim, s. 4

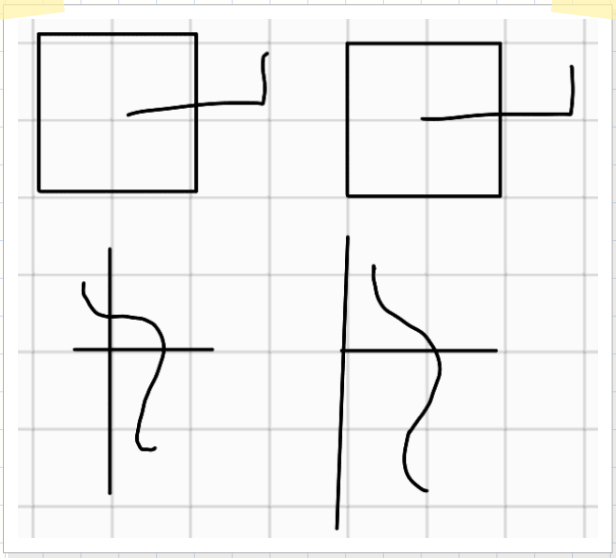
361! ihtimal

$$P(s_{t+1}, a_{t+1} \mid s_t, a_t)$$



orijinal çizim, s. 4

orj. s. 5



orijinal çizim, s. 5

Exploration
Exploitation
? Dilemma

$$V(s) = E\{G_t \mid S_t\}$$

$$G_t = \sum_{n=1}^{T} r_{t+n}$$

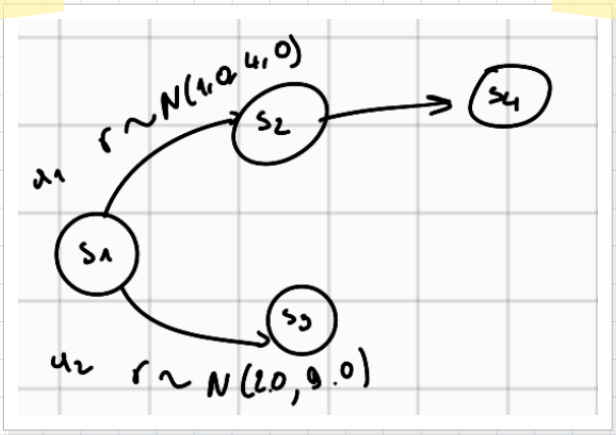
$V(s) = E\{G_t \mid S_t\}$ nasıl hesaplanır

$s_0, a_0, r_1, s_1, a_1, r_2, s_2, \ldots$

Gelecekte alacağım ? ödüllerin toplamı

$$V(s) = E\{G_t \mid S_t\}$$

$$Q(s,a) = E\{G_t \mid s_t, a\}$$



orijinal çizim, s. 5

①

$$Q(s_1, a_1) = 0$$
$$Q(s_1, a_2) = 0$$

SARSA

$$0 < \alpha < 1$$

$$Q(s_1, a_1) = Q(s_1, a_1) + \alpha\Big(\underset{\text{reward}}{r} + \overset{Q(s_2,a_2)}{V(s_2)} - Q(s_1, a_1)\Big)$$

= 2

Q learning

orj. s. 6 Policy gradient theorem

Actor critic (value, policy birleşimi)

Bunun dışında

Multi agent reinforcement learning

$$G_t = \sum_{n=1}^{T} \gamma^{n-1} r_{t+n} \quad = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3}$$

$$0 < \gamma \leq 1 \quad = r_{t+1} + \gamma G_{t+1}$$

$$V(s^t) = r_{t+1} + V(s_{t+1})$$