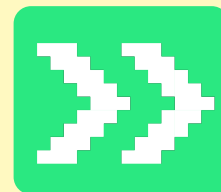


# Ранжирование через классификацию. Contrastive Learning

YOUNG & YANDEX

Радослав Нейчев  
Выпускник и преподаватель ШАД и МФТИ,  
руководитель группы ML-разработки в Яндексе,  
основатель girafe-ai



# Содержание

01

DCG

02

RankNet

03

Contrastive Learning

# DCG – Discounted Cumulative Gain

01



# DCG: Discounted Cumulative Gain

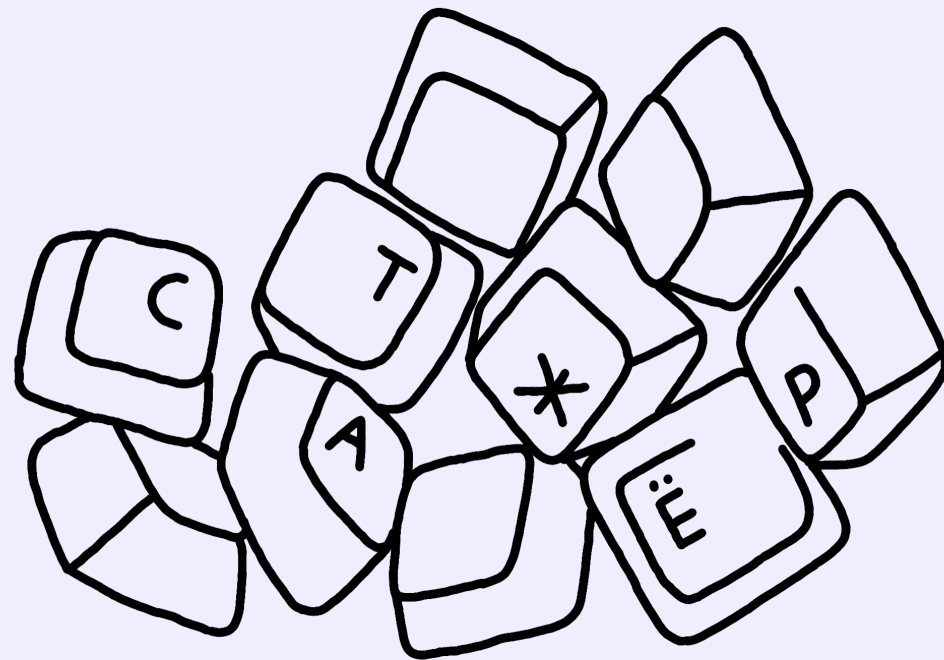
$$DCG_k = \sum_{i=1}^k \frac{rel_i}{\log_2(i + 1)}$$

или же (в других источниках)

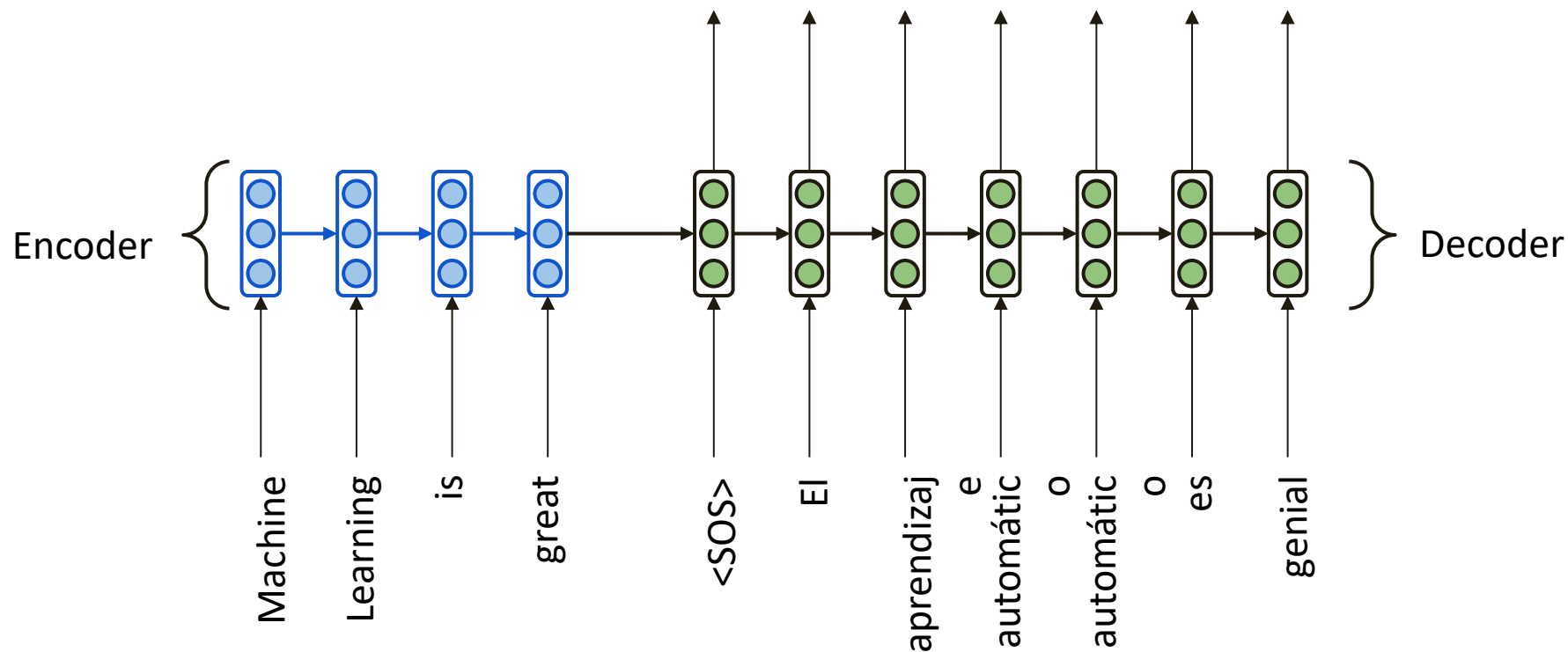
$$DCG_p = \sum_{i=1}^p \frac{2^{rel_i} - 1}{\log_2(i + 1)}$$

# RankNet

02



# Encoder-decoder архитектура



# RankNet

$$P_{ij} \equiv P(U_i \triangleright U_j) \equiv \frac{1}{1 + e^{-\sigma(s_i - s_j)}}$$

# RankNet: функция потерь

$$C = -\bar{P}_{ij} \log P_{ij} - (1 - \bar{P}_{ij}) \log(1 - P_{ij})$$

$$\bar{P}_{ij} = \frac{1}{2}(1 + S_{ij})$$

Целевая метка  $S_{ij}$  принимает следующие значения:

- 1, если документ  $i$  релевантней, чем  $j$
- -1 если документ  $j$  релевантней, чем  $i$
- 0, если их релевантность совпадает



# RankNet: функция потерь

$$C = -\bar{P}_{ij} \log P_{ij} - (1 - \bar{P}_{ij}) \log(1 - P_{ij})$$

$$w_k \rightarrow w_k - \eta \frac{\partial C}{\partial w_k} = w_k - \eta \left( \frac{\partial C}{\partial s_i} \frac{\partial s_i}{\partial w_k} + \frac{\partial C}{\partial s_j} \frac{\partial s_j}{\partial w_k} \right)$$

# RankNet: функция потерь

$$C = -\bar{P}_{ij} \log P_{ij} - (1 - \bar{P}_{ij}) \log(1 - P_{ij})$$

$$C = \frac{1}{2}(1 - S_{ij})\sigma(s_i - s_j) + \log(1 + e^{-\sigma(s_i - s_j)})$$

# RankNet: функция потерь

$$C = \frac{1}{2}(1 - S_{ij})\sigma(s_i - s_j) + \log(1 + e^{-\sigma(s_i - s_j)})$$

$$\frac{\partial C}{\partial s_i} = \sigma \left( \frac{1}{2}(1 - S_{ij}) - \frac{1}{1 + e^{\sigma(s_i - s_j)}} \right) = -\frac{\partial C}{\partial s_j}$$

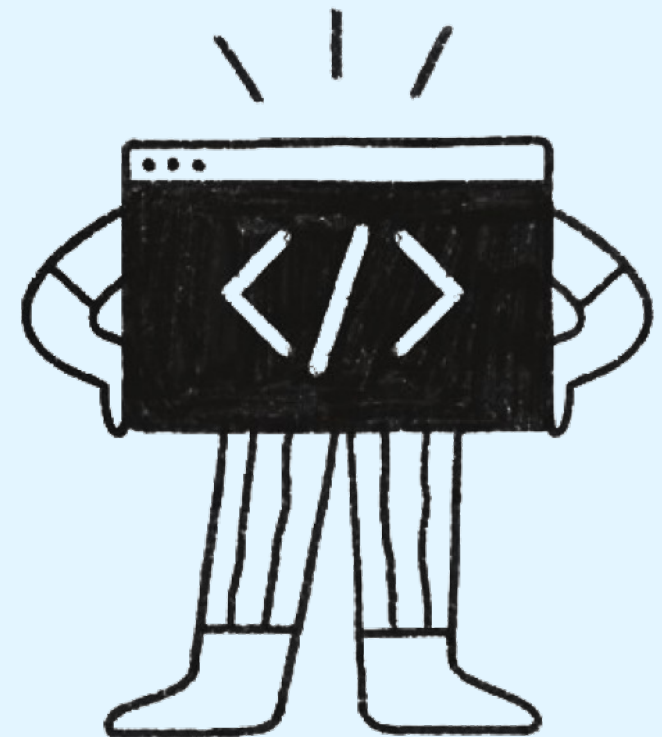
# RankNet: функция потерь

$$\frac{\partial C}{\partial s_i} = \sigma \left( \frac{1}{2}(1 - s_{ij}) - \frac{1}{1 + e^{\sigma(s_i - s_j)}} \right) = -\frac{\partial C}{\partial s_j}$$

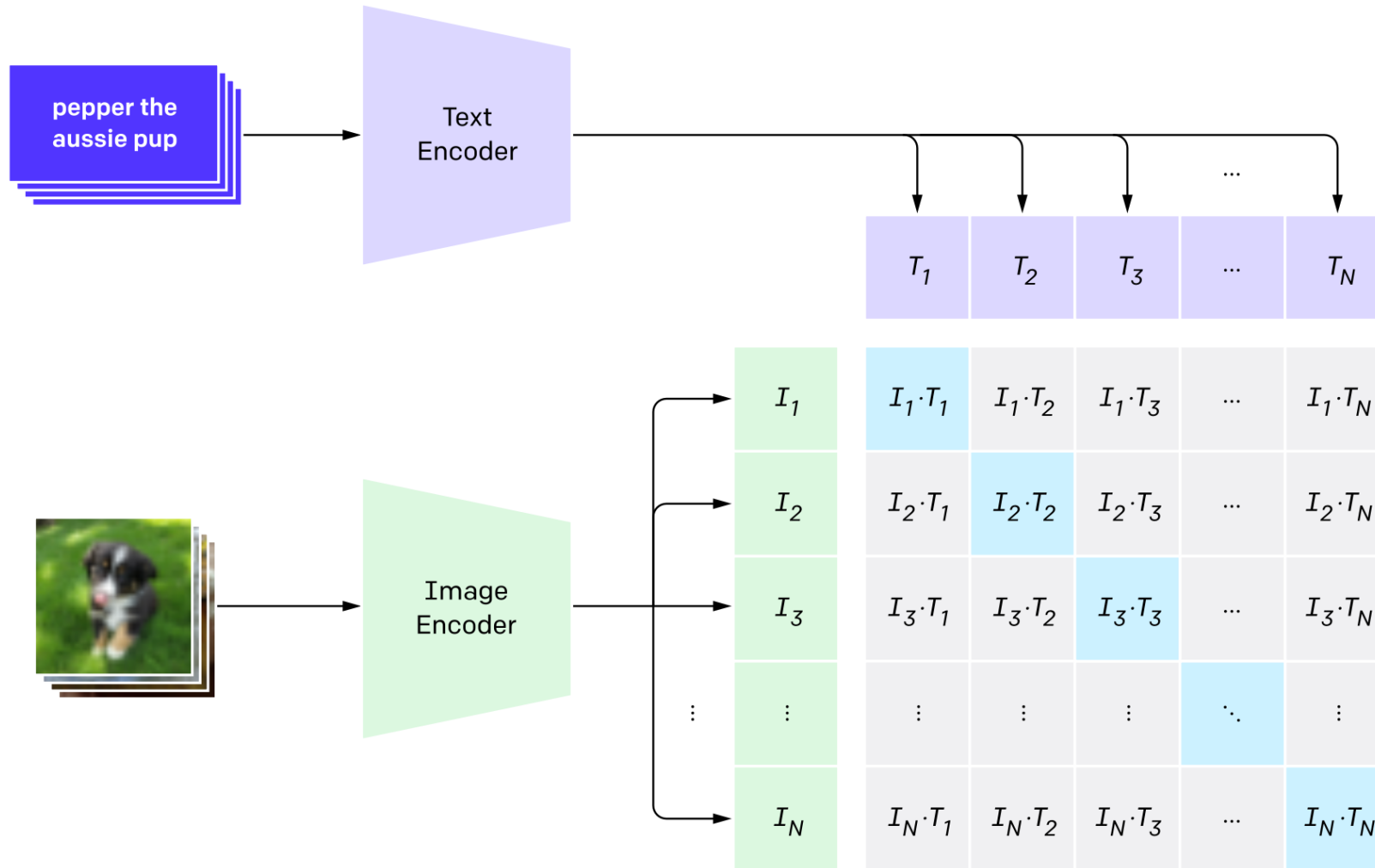
$$\begin{aligned} \frac{\partial C}{\partial w_k} &= \frac{\partial C}{\partial s_i} \frac{\partial s_i}{\partial w_k} + \frac{\partial C}{\partial s_j} \frac{\partial s_j}{\partial w_k} = \sigma \left( \frac{1}{2}(1 - s_{ij}) - \frac{1}{1 + e^{\sigma(s_i - s_j)}} \right) \left( \frac{\partial s_i}{\partial w_k} - \frac{\partial s_j}{\partial w_k} \right) \\ &= \lambda_{ij} \left( \frac{\partial s_i}{\partial w_k} - \frac{\partial s_j}{\partial w_k} \right) \end{aligned}$$

# Contrastive Learning (на примере CLIP)

03



# Contrastive Learning: CLIP



Спасибо за внимание



Y&OY