

Natural Language Processing with Transformers

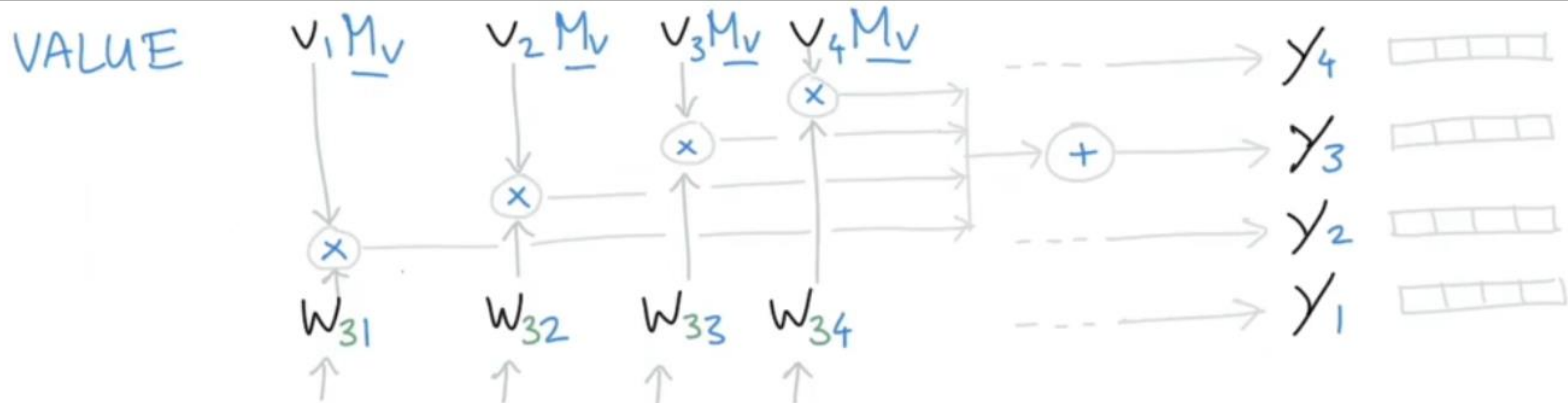
THE SELF-ATTENTION MECHANISM

- **Quiz**
- **Peer Review Breakout**
- **Evaluationsmetriken**
- **Breakout Diskussion**

QUIZ

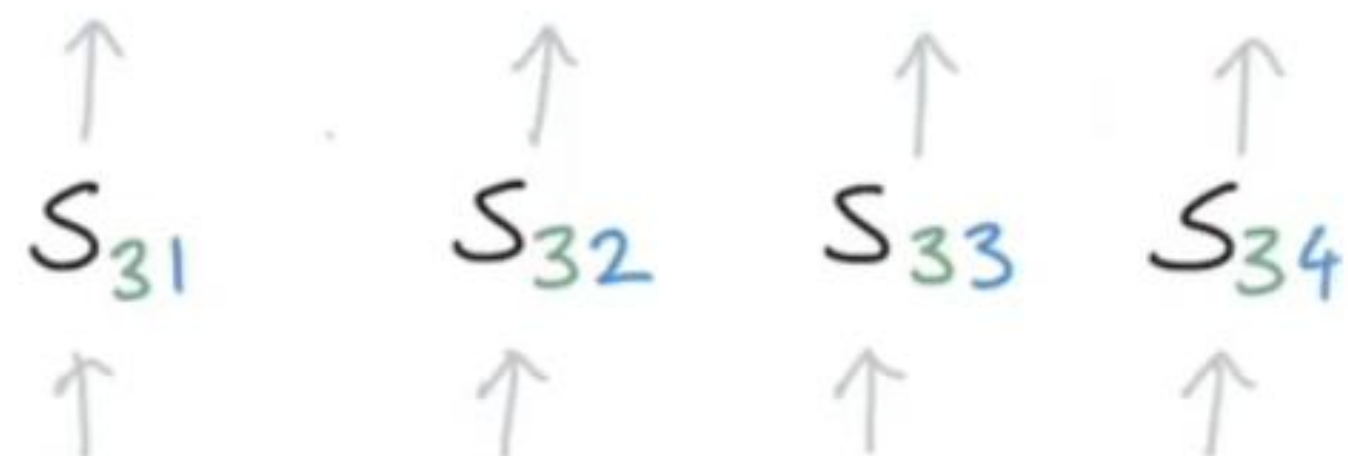


<https://forms.office.com/r/LAqEw2tLy8>



NORMALISE

$$\sum_j w_{3j} = 1$$



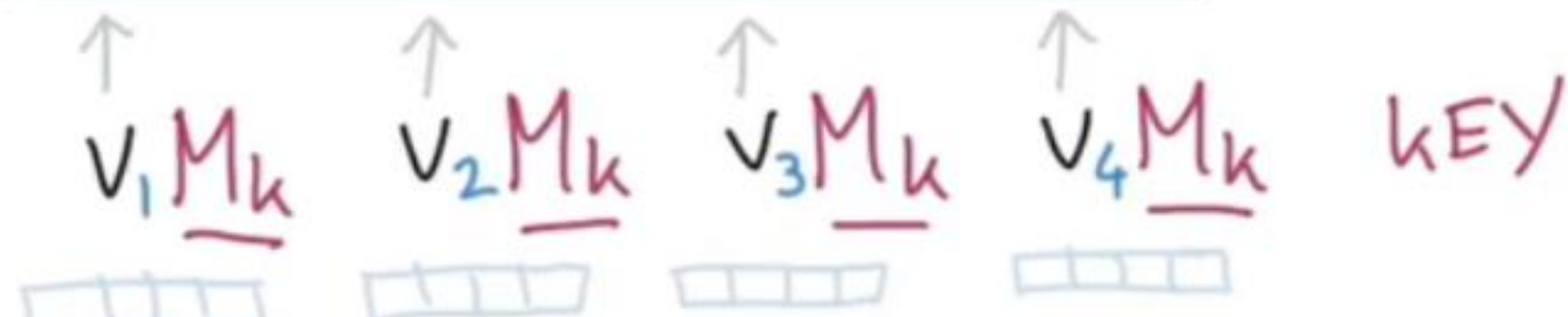
DOT PRODUCT

QUERY

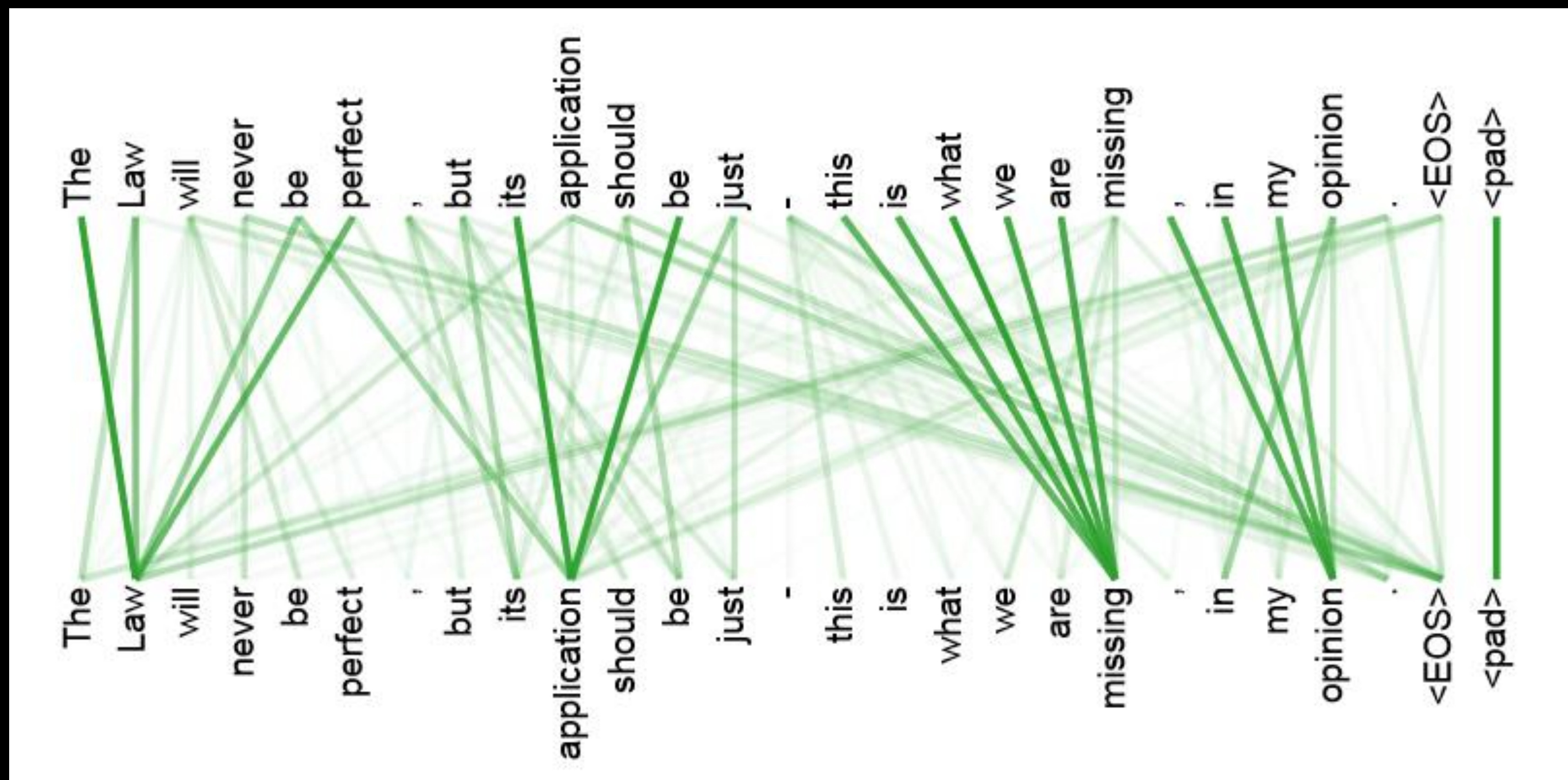
$$v_i \quad M = \begin{bmatrix} & & & \end{bmatrix}$$

$1 \times \underline{k}$ $k \times k$ $1 \times k$

$$v_3 \underline{M}_Q$$

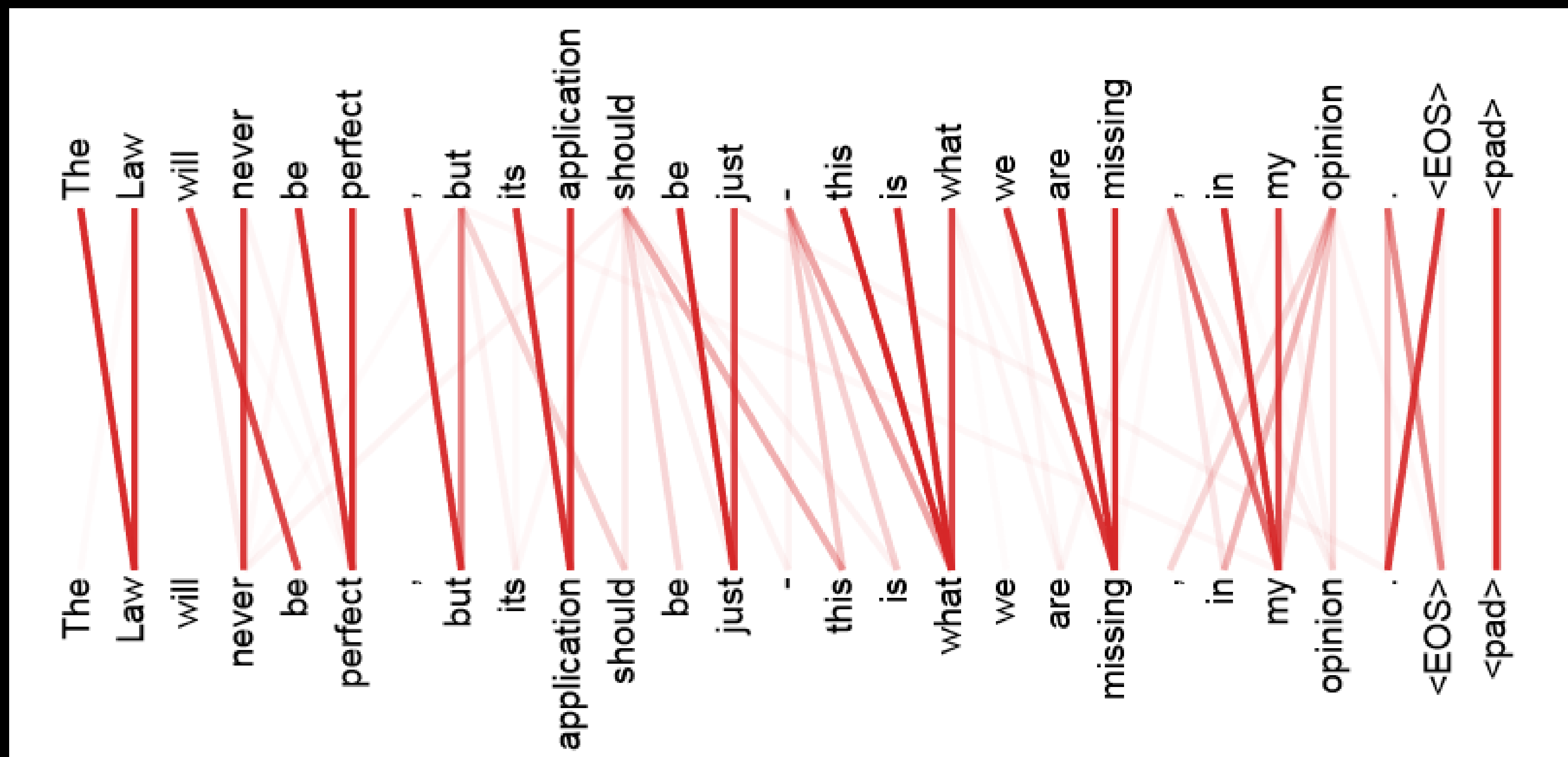


ATTENTION HEAD EXAMPLE 1



Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I.
(2017). Attention Is All You Need. *ArXiv:1706.03762 [Cs]*. <http://arxiv.org/abs/1706.03762>

ATTENTION HEAD EXAMPLE 2



Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I.
(2017). Attention Is All You Need. *ArXiv:1706.03762 [Cs]*. <http://arxiv.org/abs/1706.03762>

PEER REVIEW BREAKOUT

- **Which comparable problems and solutions did you find in the literature?**
- **What are the specific dataset characteristics (biases/limitations)?**
- **What is your baseline model?**
- **What is your evaluation metric?**

BREAKOUT ROOMS

Room 1

- **Klassifikation von Antwort-Mails hinsichtlich Höflichkeit und ggf. hinsichtlich von fachlichen Kriterien (Chris und Sabrina)**
- **Sentiment-Analyse & Themen-Tagging von Nachrichtenartikeln (Leon, Jule, Jonas, Laura, Awais)**

Room 2

- **Vorhersage der Beantwortungsschwierigkeiten von Aufgaben (Karo, Sina)**
- **SHU-T: Generierung von Antworten auf Hass-Artikel (Martin)**
- **Paraphrasing Texts (Peyman)**

HUGGING FACE COURSE

- **Token classification**
- **Fine-tuning a masked language model**
- **Translation**
- **Summarization**
- **Training a causal language model from scratch**
- **Question answering**

EVALUATION METRICS FOR LANGUAGE MODELS

- **BLUE Score**
- **ROUGE Score**
- **(Log-) Perplexity**

What is the BLEU metric?



with Lewis



ROUGE-1 SCORE

I really loved reading the Hunger Games

Machine generated summary

I loved reading the Hunger Games

Human reference summary

$$\text{ROUGE-1 recall} = \frac{\text{Num word matches}}{\text{Num words in reference}} = \frac{6}{6}$$

$$\text{ROUGE-1 precision} = \frac{\text{Num word matches}}{\text{Num words in summary}} = \frac{6}{7}$$

$$\text{ROUGE-1 F1-score} = 2 \left(\frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \right)$$

ROUGE-2 SCORE

I really

really loved

loved reading

reading the

the Hunger

Hunger Games

Generated summary
bigrams

I loved

loved reading

reading the

the Hunger

Hunger Games

Reference summary
bigrams

$$\text{ROUGE-2 recall} = \frac{\text{Num bigram matches}}{\text{Num bigrams in reference}} = \frac{4}{5}$$

$$\text{ROUGE-2 precision} = \frac{\text{Num bigram matches}}{\text{Num bigram in summary}} = \frac{4}{6}$$

ROUGE-L SCORE

I really loved reading the Hunger Games

Machine generated summary

I loved reading the Hunger Games

Human reference summary

$$\text{ROUGE-L recall} = \frac{\text{LCS}(\text{gen}, \text{ref})}{\text{Num words in reference}} = \frac{6}{6}$$

$$\text{ROUGE-L precision} = \frac{\text{LCS}(\text{gen}, \text{ref})}{\text{Num words in summary}} = \frac{6}{7}$$

LIKELIHOOD OF A SEQUENCE

$$P(X) = \prod_{i=0}^t p(x_i \mid x_{<i})$$

Hugging Face is a startup based in New York City and Paris

$p(\text{word}|\text{context})$

CROSS-ENTROPY

$$CE(X) = -\frac{1}{t} \log P(X)$$

LOG-PERPLEXITY

$$\begin{aligned} PPL(X) &= e^{CE(X)} \\ &= e^{-\frac{1}{t} \sum_{i=0}^t \log p(x_i | x_{<i})} \end{aligned}$$

BREAKOUT DISKUSSION

- **Angenommen, Ihr designt einen Prompt für ein GPT-Modell, um nicht gegenderte Texte in gegenderte Text zu überführen.**

Wie überprüft Ihr die Qualität der generierten Texte?

TODOS BIS ZUM NÄCHSTEN TERMIN

- [Video 3](#) und [Video 4](#) der Rasa Serie zu Attention und Transformers
- Weiter am Projekt arbeiten.
Versucht insbesondere unterschiedliche Evaluationsmetriken zu betrachten!