Deep Dive Sessions for Data Scientist

03 Feb 2022

Welcome!

9 00 - 9 05 Introduction

9 05 - 9 30 Information extraction

9 30 - 9 40 Breakout 1

9 40 - 9 50 Summary

9 50 - 10 00 Coffee break

10 00 - 10 40 Attention and Transformers

10 40 - 10 55 Breakout 2

10 55 - 11 00 Wrap up

AI
SWEDEN

Accelerating the use of AI for the benefit of our society, our competitiveness, and for everyone living in Sweden

Deep dive sessions

Support by targeting key roles

Provide a neutral arena where you can meet and discuss similar challenges

An outside-in perspective, inspiration & meet with other in the same situation

Data Scientist or similar function

Role description:

A Data Scientist needs to be able to:

- Autonomously and flexibly carry out advanced analytics in multiple domains and environments.
- Run advanced modeling on data in order to extract knowledge and/or predict future events.
- Autonomously create and develop concepts.
- Understand the complete analytics chain from storing, structuring, modeling, to visualizing and translating data into actionable insights.
- Review, advice and communicate in analytics use cases.

Attention, the heart of transformers, to improve Information extraction

Severine Verlinden
AI Developer, Language Technology
AI Sweden

AI
SWEDEN

Breakout session 1

Questions

- Which problems can be solved with Information Extraction in your area of work?
- Main points of the discussion

Please fill the survey:

https://docs.google.com/forms/d/e/1FAIpQLSeJcFpnNyL33o2C41fOIlqO8jrEKw1R5NqITjOJtPCK7fq58g/viewform

Break

Breakout session 2

Questions

- How do you think Attention or Transformers can be used in your area of work?
- How would you want to use language models ?

Next step

- The General Survey
- Networking on Slack

We need your feedback!

https://www.surveymonkey.com/r/MGBT6HR

AI SWEDEN


Thank you!

AI SWEDEN

Deep Dive Attention & Transformers

ai.se

youtube.com/c/aisweden

linkedin.com/company/aisweden/

What is Information Extraction?

Read more on this topic:

https://aclanthology.org/2021.findings-acl.171.pdf

Today …


Tweet
Kumail Nanjiani
@kumailn
Thread: I know there's a lot of scary stuff in the world rn, but this is something I've been thinking about that I can't get out of my head.
9:56 PM · 01 Nov 17
8,240 Retweets 17K Likes
Kumail Nanjiani @kumailn · 11h
Replying to @kumailn
As a cast member on a show about tech, our job entails visiting tech companies/ conferences etc. We meet ppl eager to show off new tech.
24 290 2,336
Kumail Nanjiani @kumailn · 11h
Often we'll see tech that is scary. I don't
Tweet your reply

Today ...

Tweet

Kumail Nanjiani
@kumailn

Thread: I know there's a lot of scary stuff in the world rn, but this is something I've been thinking about that I can't get out of my head.

9:56 PM · 01 Nov 17

8,240 Retweets 17K Likes

Kumail Nanjiani @kumailn · 11h
Replying to @kumailn
As a cast member on a show about tech, our job entails visiting tech companies/ conferences etc. We meet ppl eager to show off new tech.

24 290 2,336

Kumail Nanjiani @kumailn · 11h
Often we'll see tech that is scary. I don't

Tweet your reply

Made for minds.

TOP STORIES MEDIA CENTER TV RADIO LEARN GERMAN

GERMANY CORONAVIRUS WORLD BUSINESS SCIENCE ENVIRONMENT CULTURE

TOP STORIES / WORLD / EUROPE

EUROPE

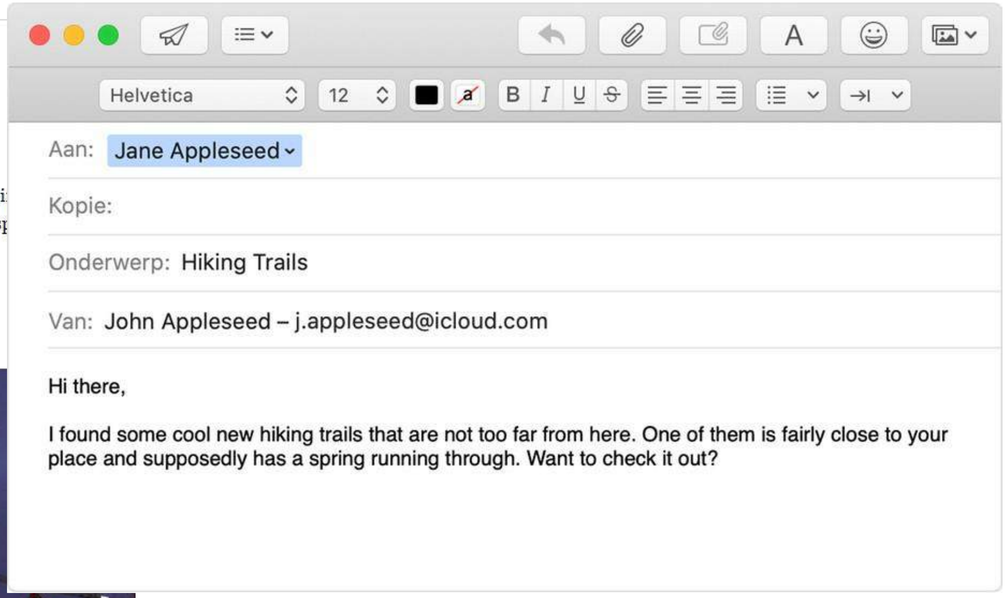
Hungary anti-LGBT+ law dispute overshadows EU summit

The European Commission and Hungary are at loggerheads over the discrimination of the LGBT+ community. The colors of the Pride rainbow were on conspicuous display in Brussels. Bernd Riegert reports.

Today ...


Tweet
Kumail Nanjiani
@kumailn
Thread: I know there's a lot of scary stuff in the world rn, but this is something I've been thinking about that I can't get out of my head.
9:56 PM · 01 Nov 17
8,240 Retweets 17K Likes
Kumail Nanjiani @kumailn · 11h
Replying to @kumailn
As a cast member on a show about tech, our job entails visiting tech companies/ conferences etc. We meet ppl eager to show off new tech.
24 290 2,336
Kumail Nanjiani @kumailn · 11h
Often we'll see tech that is scary. I don't
Tweet your reply

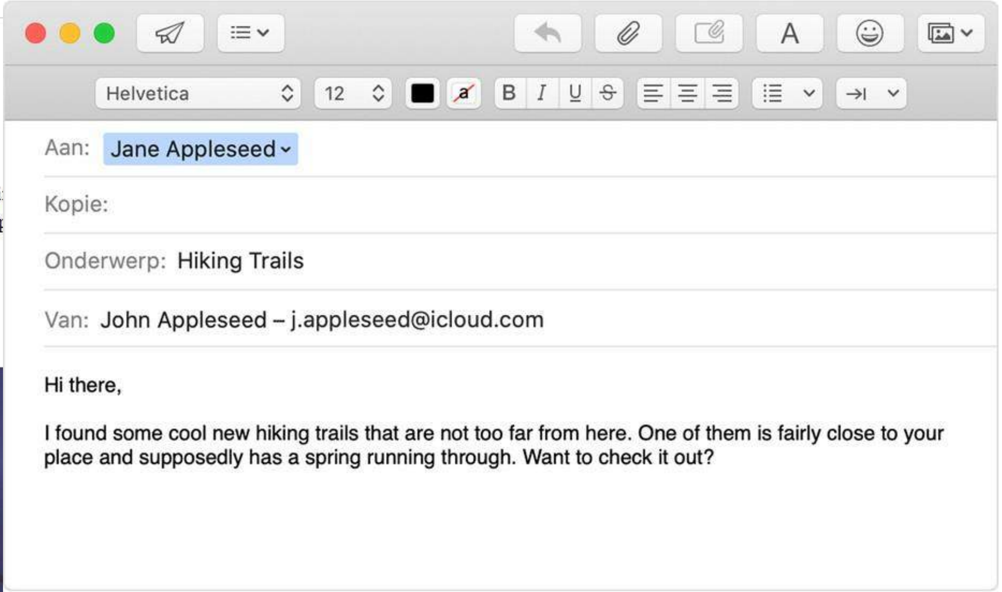

DW Made for minds.
TOP STORIES MEDIA CENTER TV RADIO LEARN GERMAN
GERMANY CORONAVIRUS WORLD BUSINESS SCIENCE ENVIRONMENT CULTURE
TOP STORIES / WORLD / EUROPE
EUROPE
Hungary anti-LGBT+ law dispute overshadows EU summit
The European Commission and Hungary are at loggerheads over the discrimi
LGBT+ community. The colors of the Pride rainbow were on conspicuous disp
Brussels. Bernd Riegert reports.


Helvetica 12
Aan: Jane Appleseed
Kopie:
Onderwerp: Hiking Trails
Van: John Appleseed – j.appleseed@icloud.com
Hi there,
I found some cool new hiking trails that are not too far from here. One of them is fairly close to your place and supposedly has a spring running through. Want to check it out?

Today ...


Tweet
Kumail Nanjiani
@kumailn
Thread: I know there's a lot of scary stuff in the world rn, but this is something I've been thinking about that I can't get out of my head.
9:56 PM - 01 Nov 17
8,240 Retweets 17K Likes
Kumail Nanjiani @kumailn · 11h
Replying to @kumailn
As a cast member on a show about tech, our job entails visiting tech companies/ conferences etc. We meet ppl eager to show off new tech.
24 290 2,336
Kumail Nanjiani @kumailn · 11h
Often we'll see tech that is scary. I don't
Tweet your reply

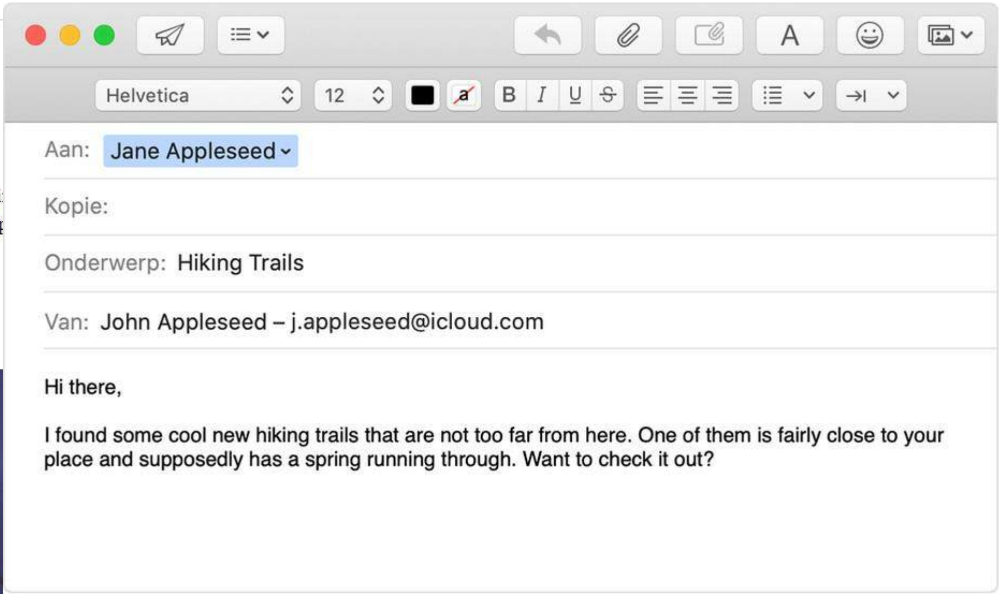

DW Made for minds.
TOP STORIES MEDIA CENTER TV RADIO LEARN GERMAN
GERMANY CORONAVIRUS WORLD BUSINESS SCIENCE ENVIRONMENT CULTURE
TOP STORIES / WORLD / EUROPE
EUROPE
Hungary anti-LGBT+ law dispute overshadows EU summit
The European Commission and Hungary are at loggerheads over the discrimi
LGBT+ community. The colors of the Pride rainbow were on conspicuous disp
Brussels. Bernd Riegert reports.


Helvetica 12
Aan: Jane Appleseed
Kopie:
Onderwerp: Hiking Trails
Van: John Appleseed – j.appleseed@icloud.com
Hi there,
I found some cool new hiking trails that are not too far from here. One of them is fairly close to your place and supposedly has a spring running through. Want to check it out?

Today ...

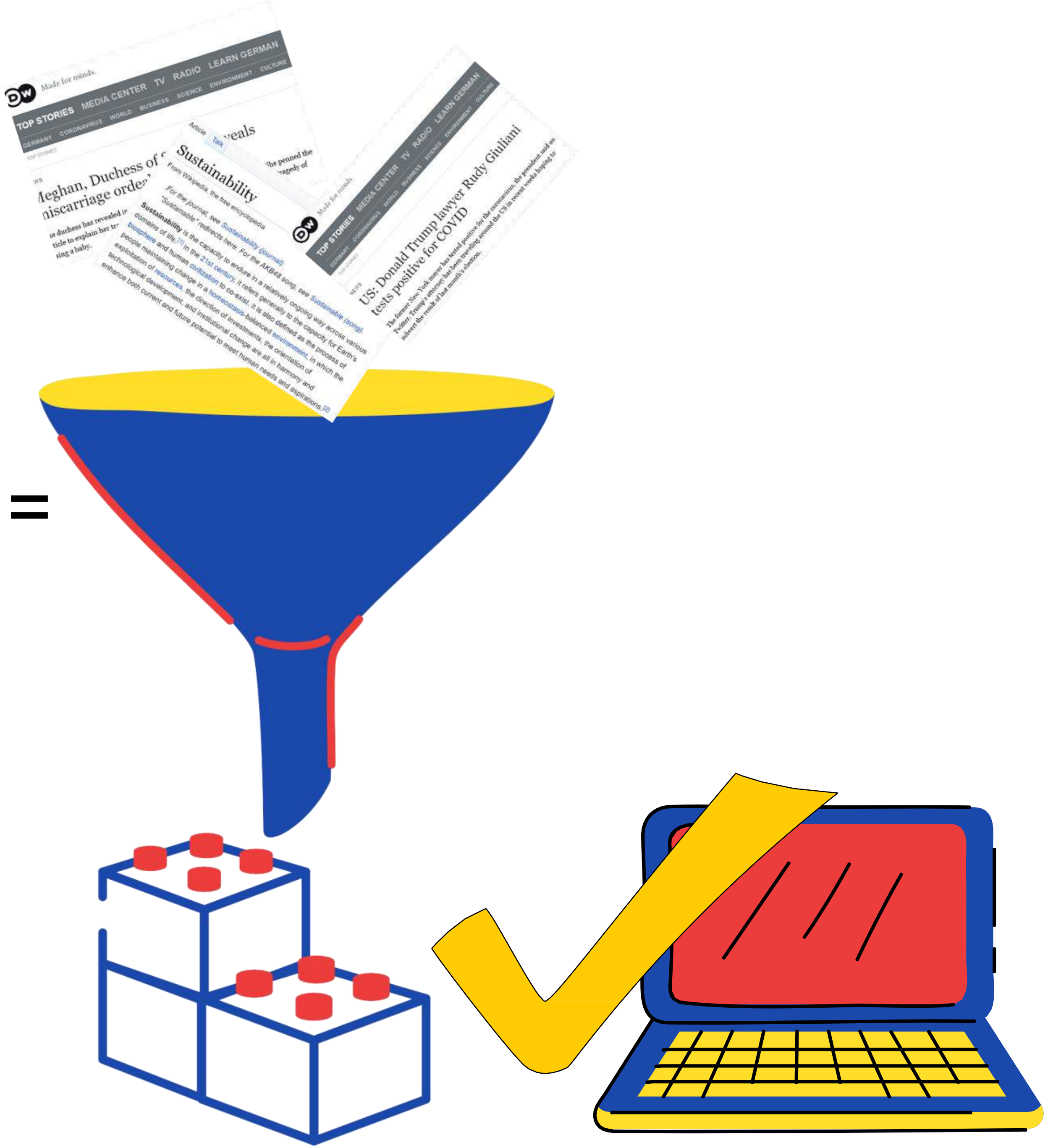
Tweet
Kumail Nanjiani
@kumailn
Thread: I know there's a lot of scary stuff in the world rn, but this is something I've been thinking about that I can't get out of my head.
9:56 PM · 01 Nov 17
8,240 Retweets 17K Likes
Kumail Nanjiani @kumailn · 11h
Replying to @kumailn
As a cast member on a show about tech, our job entails visiting tech companies/ conferences etc. We meet ppl eager to show off new tech.
24 290 2,336
Kumail Nanjiani @kumailn · 11h
Often we'll see tech that is scary. I don't
Tweet your reply

DW Made for minds.
TOP STORIES MEDIA CENTER TV RADIO LEARN GERMAN
GERMANY CORONAVIRUS WORLD BUSINESS SCIENCE ENVIRONMENT CULTURE
TOP STORIES / WORLD / EUROPE
EUROPE
Hungary anti-LGBT+ law dispute overshadows EU summit
The European Commission and Hungary are at loggerheads over the discrimi
LGBT+ community. The colors of the Pride rainbow were on conspicuous disp
Brussels. Bernd Riegert reports.



Helvetica 12
Aan: Jane Appleseed
Kopie:
Onderwerp: Hiking Trails
Van: John Appleseed – j.appleseed@icloud.com
Hi there,
I found some cool new hiking trails that are not too far from here. One of them is fairly close to your place and supposedly has a spring running through. Want to check it out?

Information Extraction =



Concrete example of information extraction

"We are very excited for Harry and Meghan. It has been wonderful getting to know Meghan and to see how happy she and Harry are together," Clarence House said in a tweet.

• Named Entity recognition	Mentions: Harry, Meghan, Clarence House...
• Coreference Resolution	Cluster: (Harry, Harry), (Meghan , Meghan), ...
• Relation Extraction	Relation: Meghan-> in relation with ->Harry

Concrete example of information extraction

"We are very excited for Harry and Meghan. It has been wonderful getting to know Meghan and to see how happy she and Harry are together," Clarence House said in a tweet.

- Named Entity recognition — Mentions: Harry, Meghan, Clarence House...
- Coreference Resolution — Cluster: (Harry, Harry), (Meghan , Meghan), ...
- Relation Extraction — Relation: Meghan–> in relation with –>Harry

Concrete example of information extraction

"We are very excited for Harry and **Meghan**. It has been wonderful getting to know **Meghan** and to see how happy she and Harry are together," Clarence House said in a tweet.

- Named Entity recognition
- **Coreference Resolution**
- Relation Extraction

Mentions: Harry, Meghan, Clarence House...

Cluster: (Harry, Harry), (Meghan , Meghan), ...

Relation: Meghan–> in relation with –>Harry

Concrete example of information extraction

"We are very excited for Harry and **Meghan**. It has been wonderful getting to know **Meghan** and to see how happy she and Harry are together," Clarence House said in a tweet.

- Named Entity recognition
- **Coreference Resolution**
- Relation Extraction

Mentions: Harry, Meghan, Clarence House...

Cluster: (Harry, Harry), (Meghan , Meghan), ...

Relation: Meghan–> in relation with **?** –>Harry

Concrete example of information extraction

"We are very excited for Harry and **Meghan**. It has been wonderful getting to know **Meghan** and to see how happy she and Harry are together," Clarence House said in a tweet.

Concrete example of information extraction

"We are very excited for Harry and **Meghan**. It has been wonderful getting to know **Meghan** and to see how happy she and Harry are together," Clarence House said in a tweet.

Concrete example of information extraction

"We are very excited for Harry and **Meghan**. It has been wonderful getting to know **Meghan** and to see how happy she and Harry are together," Clarence House said in a tweet.

=> Wikipedia, Wikidata

Problem Statement

Improvement of the current IE algorithm

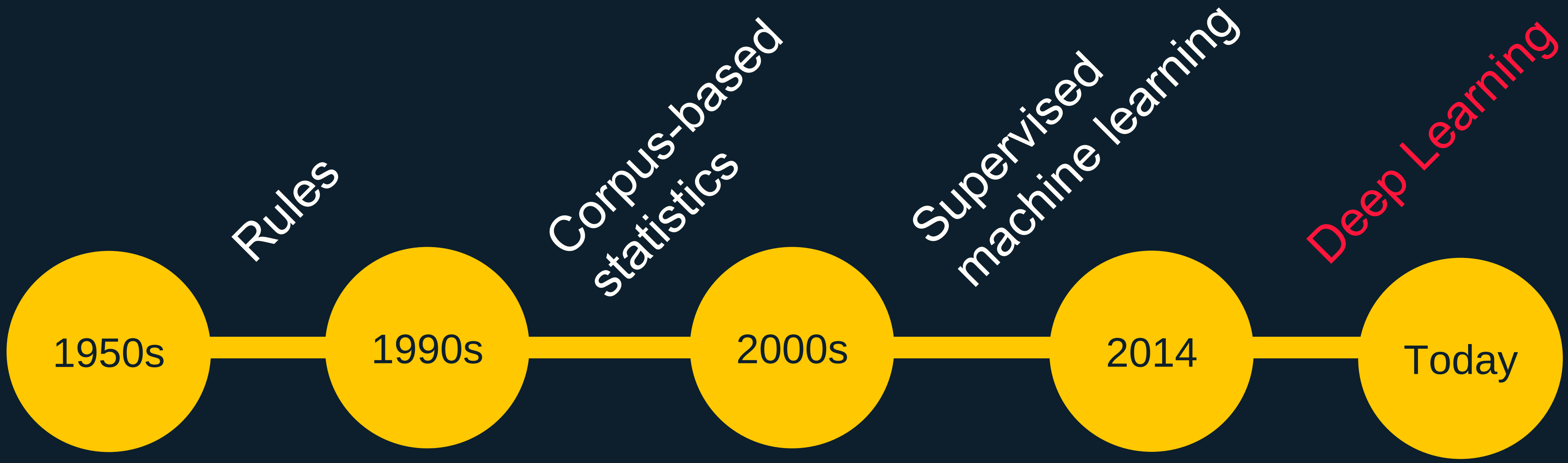
Including external knowledge of a knowledge base.

- Wikipedia
- Wikidata

Methods

Part 02

Timeline


1950s
Rules
1990s
Corpus-based statistics
2000s
Supervised machine learning
2014
Deep Learning
Today

WORD EMBEDDING

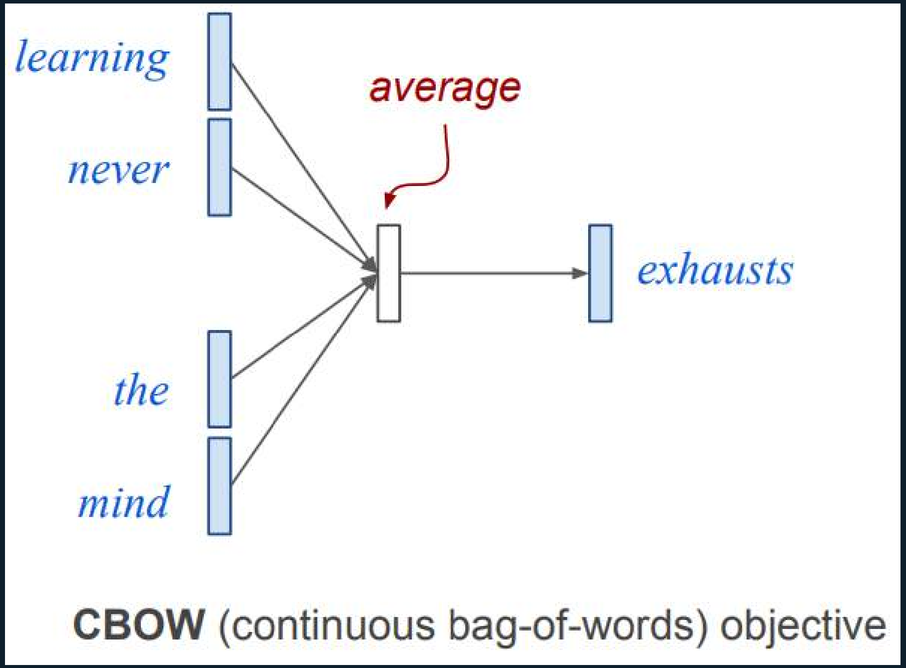
Woman	
Queen	0
Apple	0
Man	0
King	1
Orange	0
	0

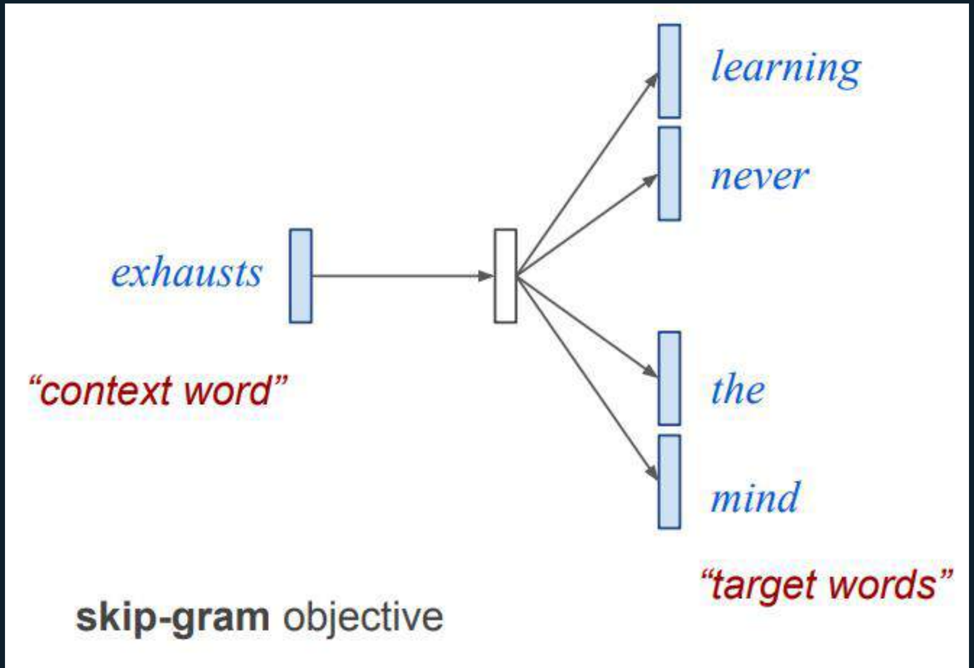
|V| X 1

	Gender	Royal	Age
Woman	1	0.02	0.02
Queen	0.98	0.95	0.69
Apple	−0.01	0.00	0.03
Man	−1	0.01	0.03
King	−0.95	0.93	0.70
Orange	0.00	0.01	−0.02

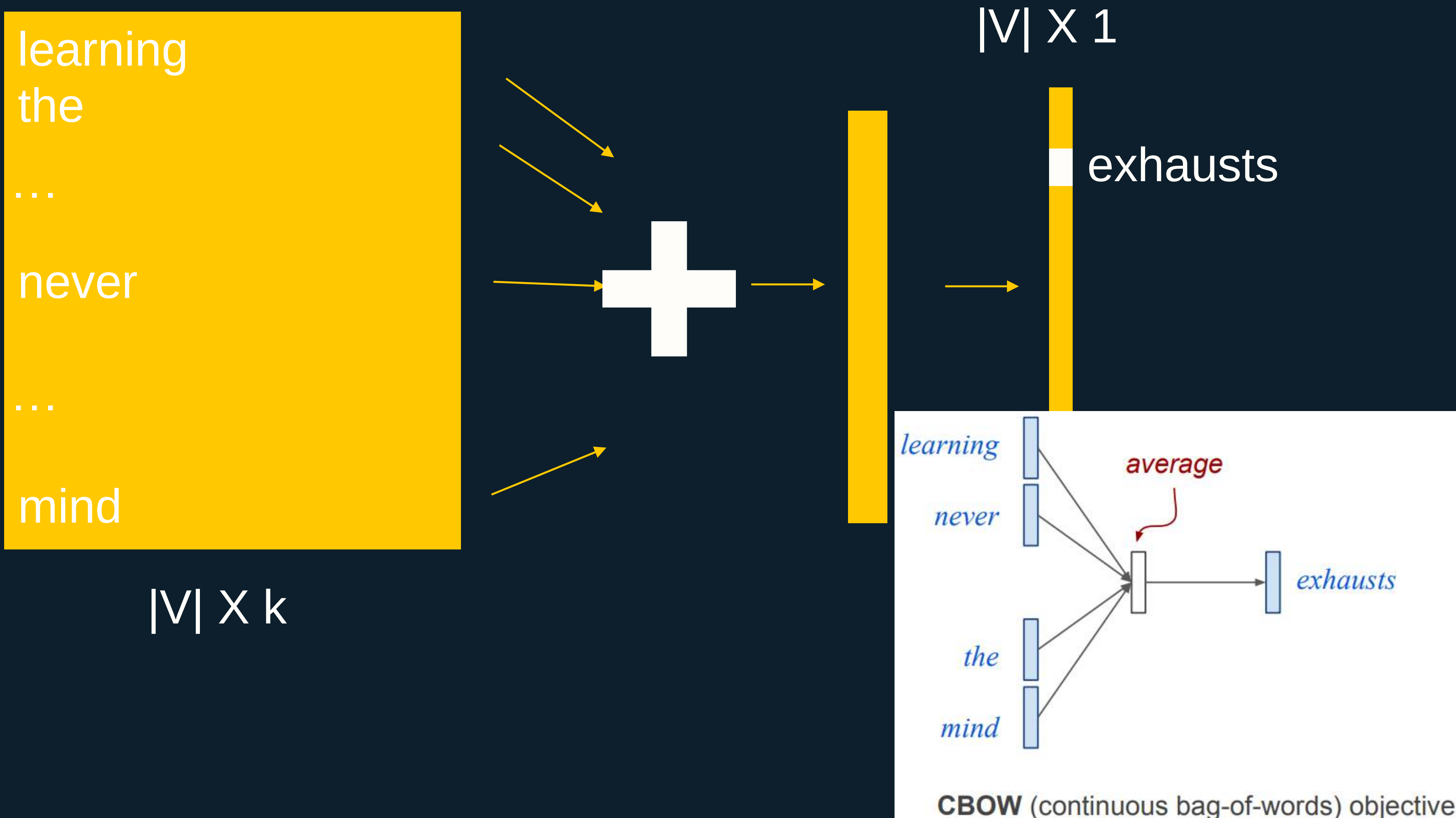
|V| X k

WORD EMBEDDING





WORD EMBEDDING



How? (*)

- NER
- RE
- CR

=>End-to-end

=>Jointly

=>Document-level

(*)Klim Zaporojets, Johannes Deleu, Chris Develder, and Thomas Demeester. 2021.
DWIE: An entity-centric dataset for multi-task document-level information extraction.
Information Processing & Management,58(4):102563

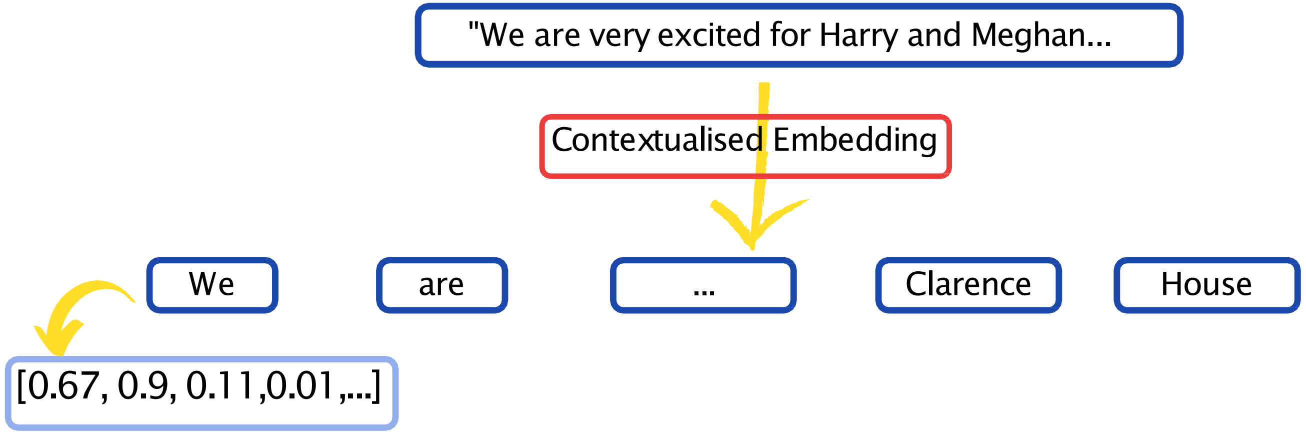
How? (*)

- NER
- RE
- CR

=>End-to-end

=>Jointly

=>Document-level



(*)Klim Zaporojets, Johannes Deleu, Chris Develder, and Thomas Demeester. 2021. DWIE: An entity-centric dataset for multi-task document-level information extraction. Information Processing & Management,58(4):102563

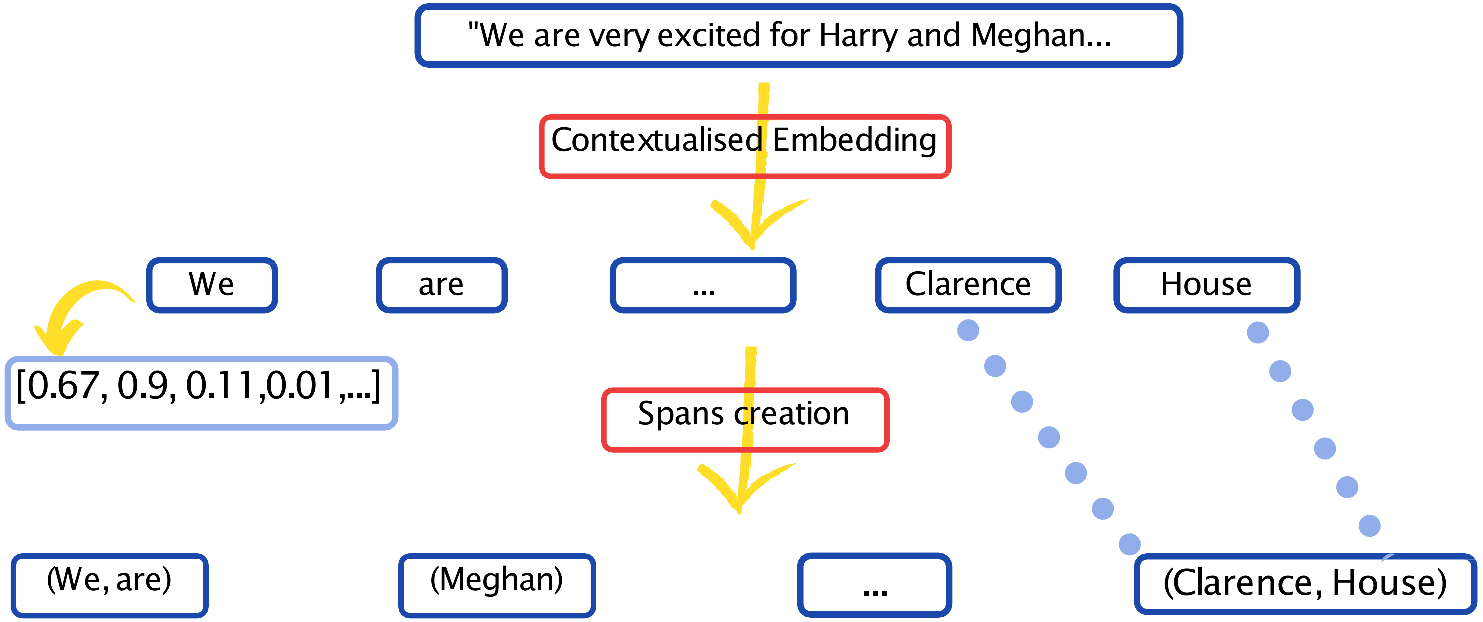
How? (*)

- NER
- RE
- CR

=>End-to-end

=>Jointly

=>Document-level

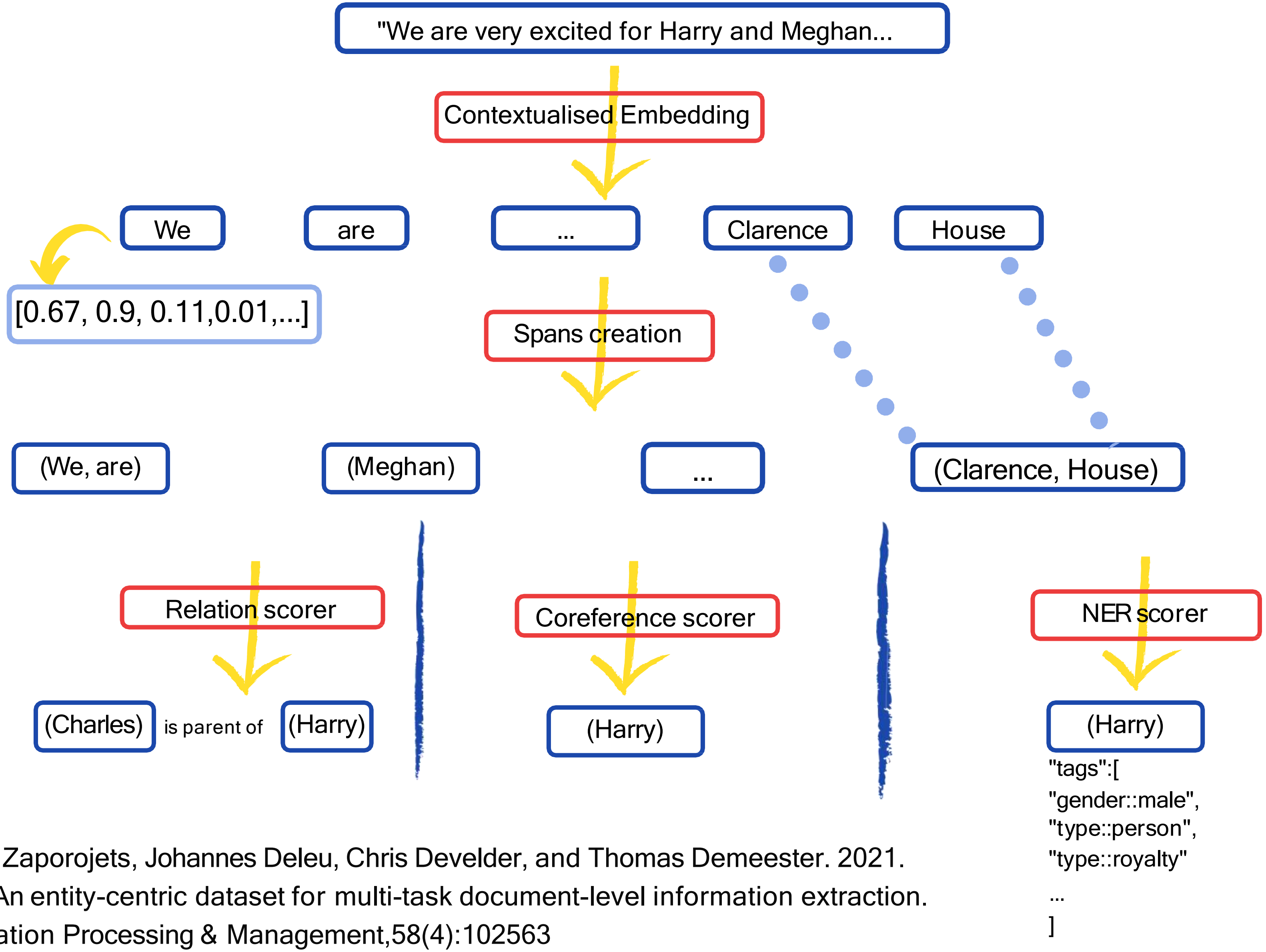


(*)Klim Zaporojets, Johannes Deleu, Chris Develder, and Thomas Demeester. 2021. DWIE: An entity-centric dataset for multi-task document-level information extraction. Information Processing & Management,58(4):102563

How? (*)

- NER
- RE
- CR

=>End-to-end
=>Jointly
=>Document-level



(*)Klim Zaporojets, Johannes Deleu, Chris Develder, and Thomas Demeester. 2021. DWIE: An entity-centric dataset for multi-task document-level information extraction. Information Processing & Management,58(4):102563

How? (*)

- NER
- RE
- CR

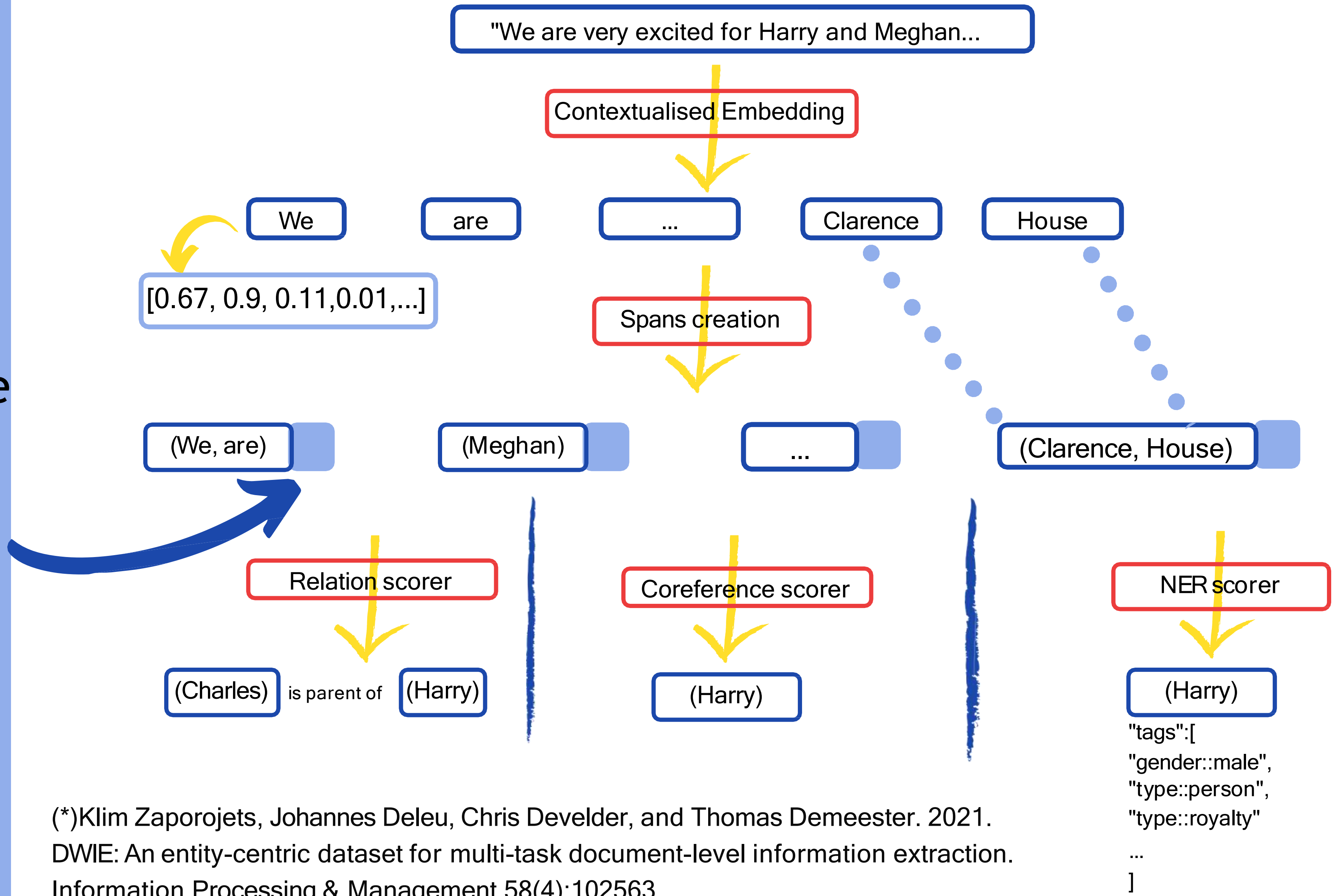
=>End-to-end

=>Jointly

=>Document-level

External Knowledge

- Wikipedia
- Wikidata



(*)Klim Zaporojets, Johannes Deleu, Chris Develder, and Thomas Demeester. 2021. DWIE: An entity-centric dataset for multi-task document-level information extraction. Information Processing & Management,58(4):102563

External Knowledge

Article Talk Read View source

Meghan, Duchess of Sussex

From Wikipedia, the free encyclopedia

Meghan, Duchess of Sussex (/ˈmɛɡən/; born **Rachel Meghan Markle**, August 4, 1981), is an American member of the British royal family and a former actress.

Markle was born and raised in Los Angeles, California. Her acting career began while she was studying at Northwestern University. She attributed early career difficulties to her biracial heritage. Her most significant acting role is that of Rachel Zane in the American legal drama *Suits*, in which she starred for seven seasons (2011–2018). She simultaneously profited from a strong social media presence, including a lifestyle blog, *The Tig* (2014–2017). Through *The Tig* she gained recognition for her fashion sense, which led to creating and releasing two lines of clothing in 2015–2016. During her acting career, Markle became involved in charity work, focusing primarily on women's issues and social justice.

External Knowledge

External Knowledge

(Meghan)

Candidates	Prior
1. "Meghan_Trainor"	0.4
2. "Meghan_McCain"	0.3
3. "Meghan,_Duchess_of_Sussex"	0.2
4. "Megan"	0.1
5. "Meghan_Allen"	0.0


Article Talk Read View source

Meghan, Duchess of Sussex

From Wikipedia, the free encyclopedia

Meghan, Duchess of Sussex (/ˈmɛɡən/; born **Rachel Meghan Markle**, August 4, 1981), is an American member of the British royal family and a former actress.

Markle was born and raised in Los Angeles, California. Her acting career began while she was studying at Northwestern University. She attributed early career difficulties to her biracial heritage. Her most significant acting role is that of Rachel Zane in the American legal drama *Suits*, in which she starred for seven seasons (2011–2018). She simultaneously profited from a strong social media presence, including a lifestyle blog, *The Tig* (2014–2017). Through *The Tig* she gaine

creating and releasing two

career, Markle became invo

women's issues and social

Article Talk Read Edit View history Search Wikipedia

Multiracial people

From Wikipedia, the free encyclopedia

(Redirected from Biracial)

"Mixed race" redirects here. For the album by Tricky, see Mixed Race (album). For more information, see Multiracialism.

External Knowledge

(Meghan)

Candidates	Prior
1. "Meghan_Trainor"	0.4
2. "Meghan_McCain"	0.3
3. "Meghan,_Duchess_of_Sussex"	0.2
4. "Megan"	0.1
5. "Meghan_Allen"	0.0

KB-text: [0.67, 0.9, 0.11,0.01,...]


Article Talk Read View source

Meghan, Duchess of Sussex

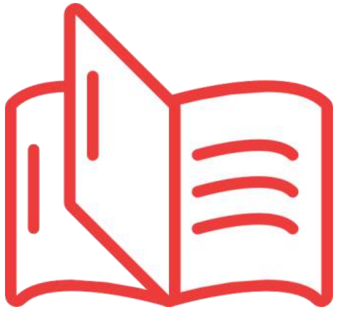
From Wikipedia, the free encyclopedia

Meghan, Duchess of Sussex (/ˈmɛɡən/; born **Rachel Meghan Markle**, August 4, 1981), is an American member of the British royal family and a former actress.

Markle was born and raised in Los Angeles, California. Her acting career began while she was studying at Northwestern University. She attributed early career difficulties to her biracial heritage. Her most significant acting role is that of Rachel Zane in the American legal drama *Suits*, in which she starred for seven seasons (2011–2018). She simultaneously profited from a strong social media presence, including a lifestyle blog, *The Tig* (2014–2017). Through *The Tig* she gained recognition for her fashion sense, which led to creating and releasing two lines of clothing in 2015–2016. During her acting career, Markle became involved in charity work, focusing primarily on women's issues and social justice.

Different configurations

(Meghan)

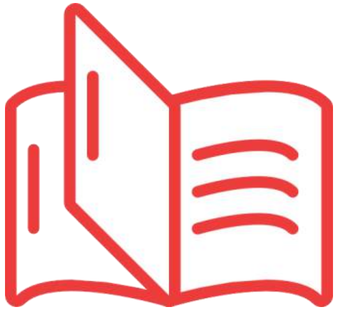


Candidates

1. "Meghan_Trainor"
2. "Meghan_McCain"
3. "Meghan,_Duchess_of_Sussex"
4. "Megan"
5. "Meghan_Allen"

Different configurations

(Meghan)



Candidates

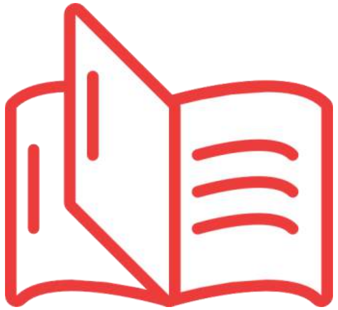
1. "Meghan_Trainor"
2. "Meghan_McCain"
3. "Meghan,_Duchess_of_Sussex"
4. "Megan"
5. "Meghan_Allen"

1weigthed entity embedding

(Meghan)

Different configurations

(Meghan)



Candidates

1. "Meghan_Trainor"
2. "Meghan_McCain"
3. "Meghan,_Duchess_of_Sussex"
4. "Megan"
5. "Meghan_Allen"

α

1weigthed entity embedding

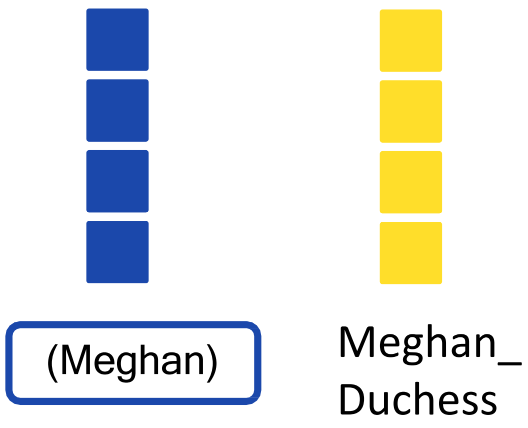
(Meghan)

Oracle

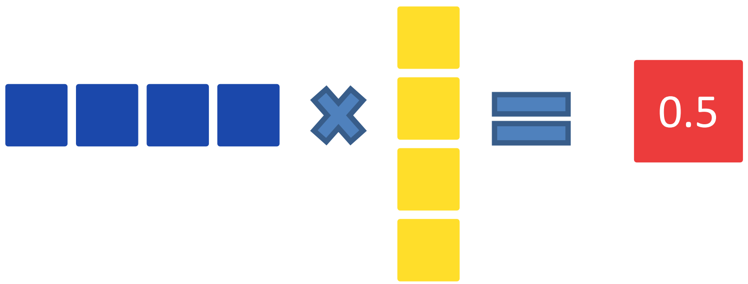
Attention

Prior

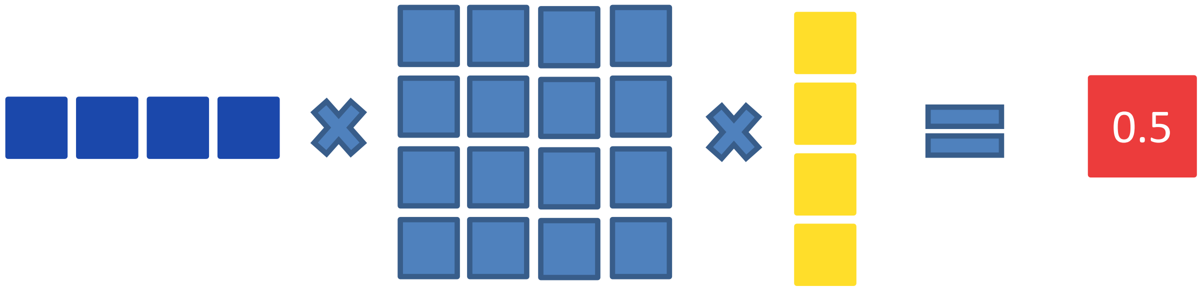
Attention



- Basic dot-product attention:



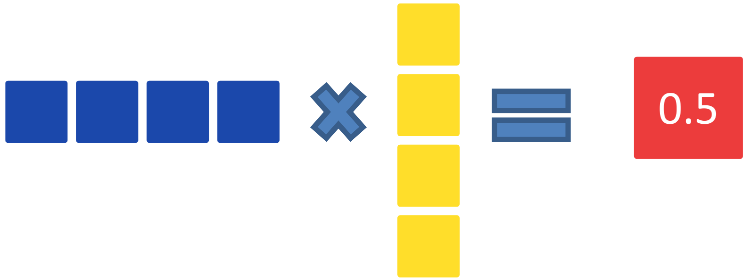
- Multiplicative attention:

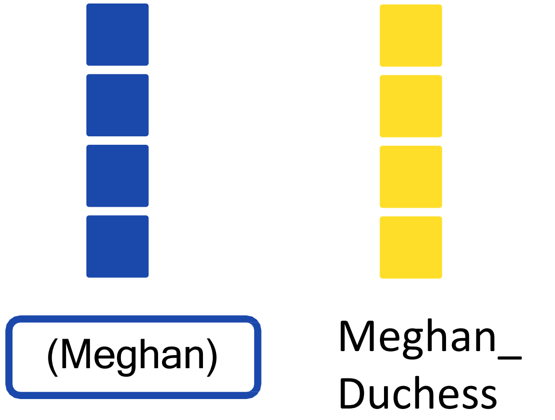


- Additive attention:

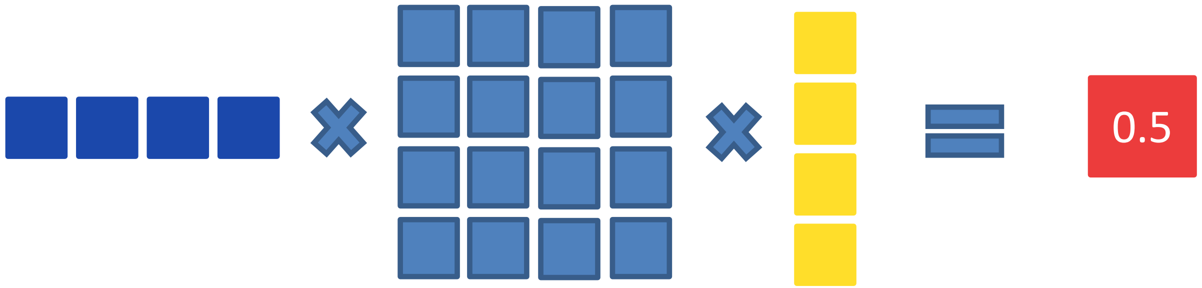
Attention

- Basic dot-product attention:





- Multiplicative attention:



- Additive attention:

Different configurations α?

Oracle

"Meghan_Trainor"

"Meghan_McCain"

"Meghan,_Duchess"

"Meghan"

"Meghan_Allen"

(Meghan)

Different configurations α?

"Meghan_Trainor"
"Meghan_McCain"
"Meghan,_Duchess"
"Meghan"
"Meghan_Allen"

(Meghan)

Oracle

Prior

0.4 0.3 0.2 0.1 0.0

Different configurations α?

(Meghan)

	"Meghan_Trainor"	"Meghan_McCain"	"Meghan,_Duchess"	"Meghan"	"Meghan_Allen"
Oracle					
Prior	0.4	0.3	0.2	0.1	0.0
Attention	0.1	0.3	0.5	0.0	0.0

Results

Part 03

Q3: How does external knowledge influence the prediction results?

CR

THE CLUSTERING
BECOMES MORE FINE
GRAINED

Q3: How does external knowledge influence the prediction results?

CR

THE CLUSTERING
BECOMES MORE FINE
GRAINED

Google VS Google street view

Q3: How does external knowledge influence the prediction results?

CR

THE CLUSTERING BECOMES MORE FINE GRAINED

NER

IMPROVES THE PREDICTION OF RARE ENTITY TYPES

Google VS Google street view

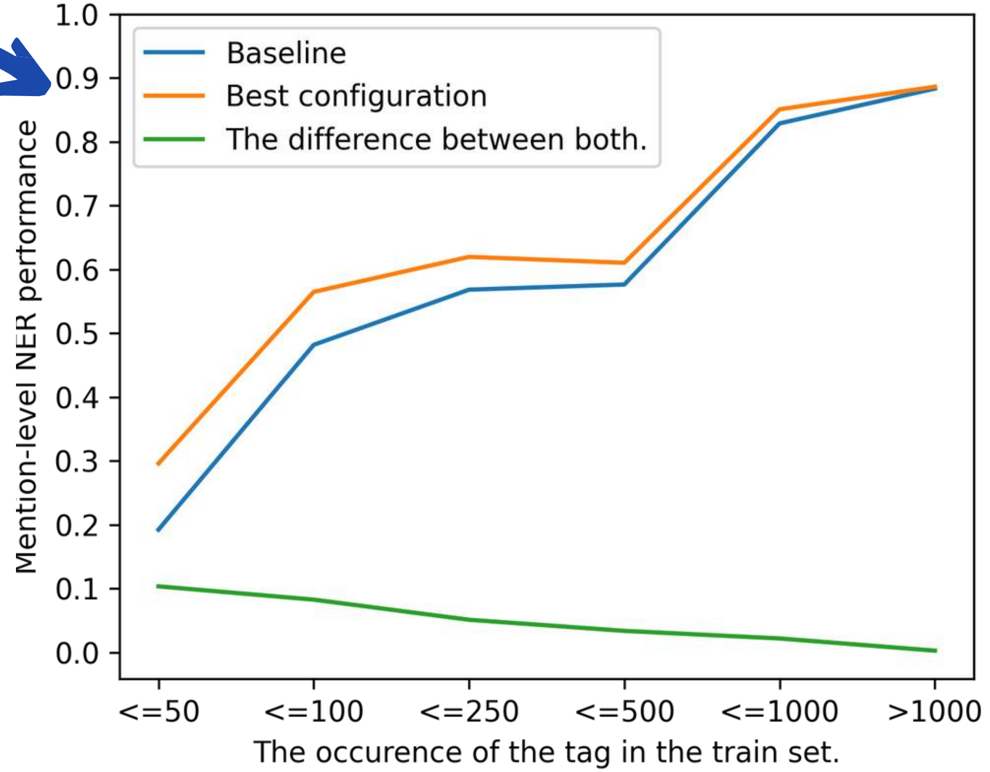
Q3: How does external knowledge influence the prediction results?

CR

THE CLUSTERING BECOMES MORE FINE GRAINED

NER

IMPROVES THE PREDICTION OF RARE ENTITY TYPES



Google VS Google street view

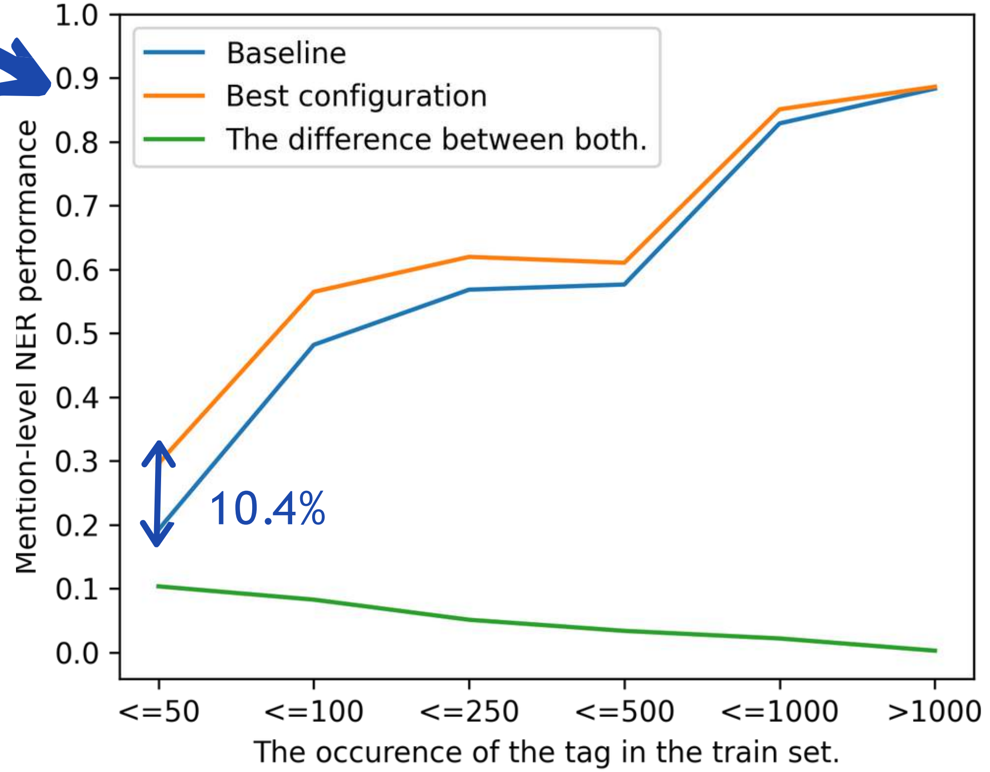
Q3: How does external knowledge influence the prediction results?

CR

THE CLUSTERING BECOMES MORE FINE GRAINED

NER

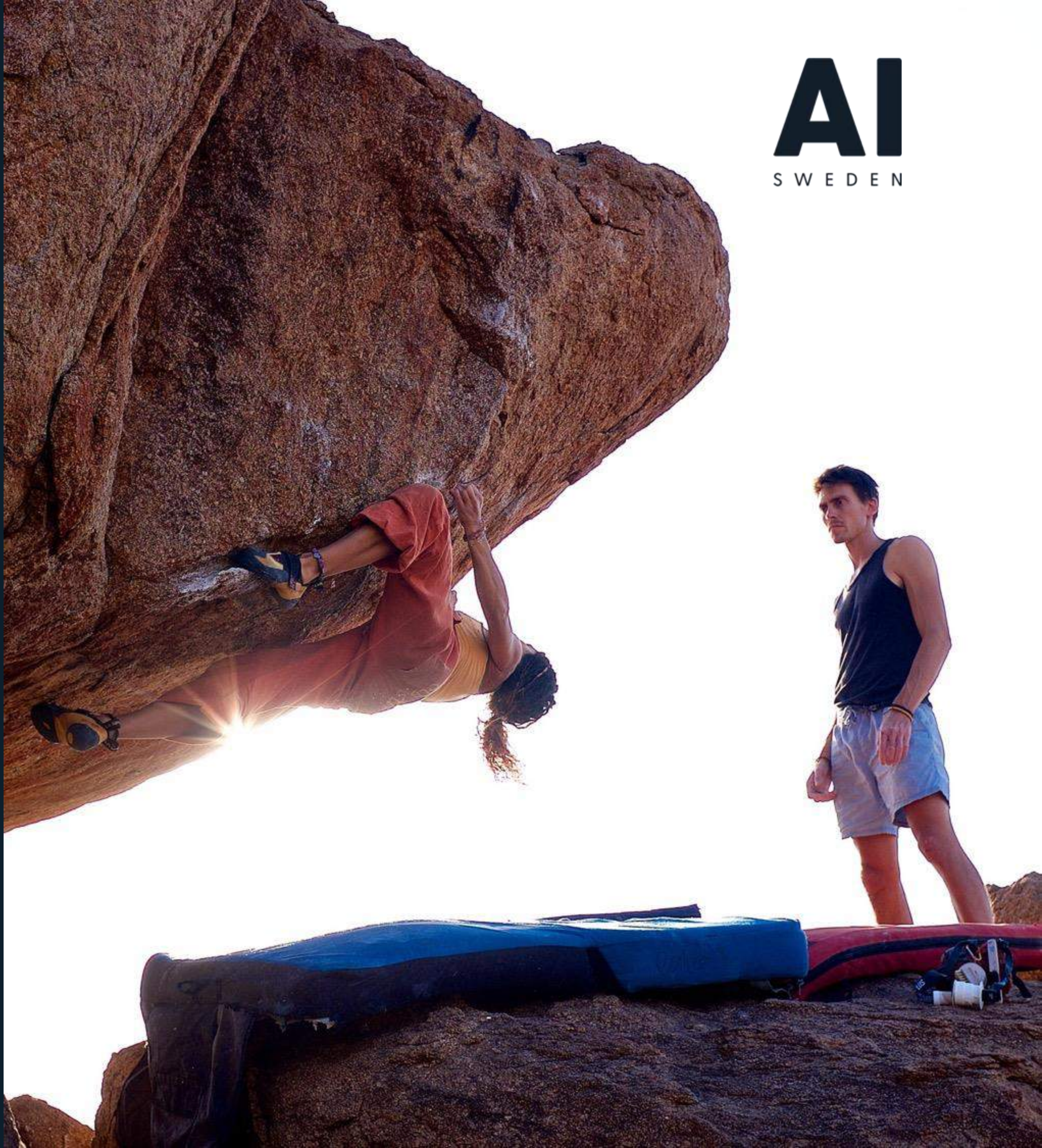
IMPROVES THE PREDICTION OF RARE ENTITY TYPES



Google VS Google street view

Questions

1. Which problems can be solved with Information Extraction in your current work field?



Transformers

Transformers =

Language Models

Attention:

-long range dependencies

- Omnidirectional

AI
SWEDEN

Transformers are

Language Models

Language Model

Language Modeling = Models the probability of text

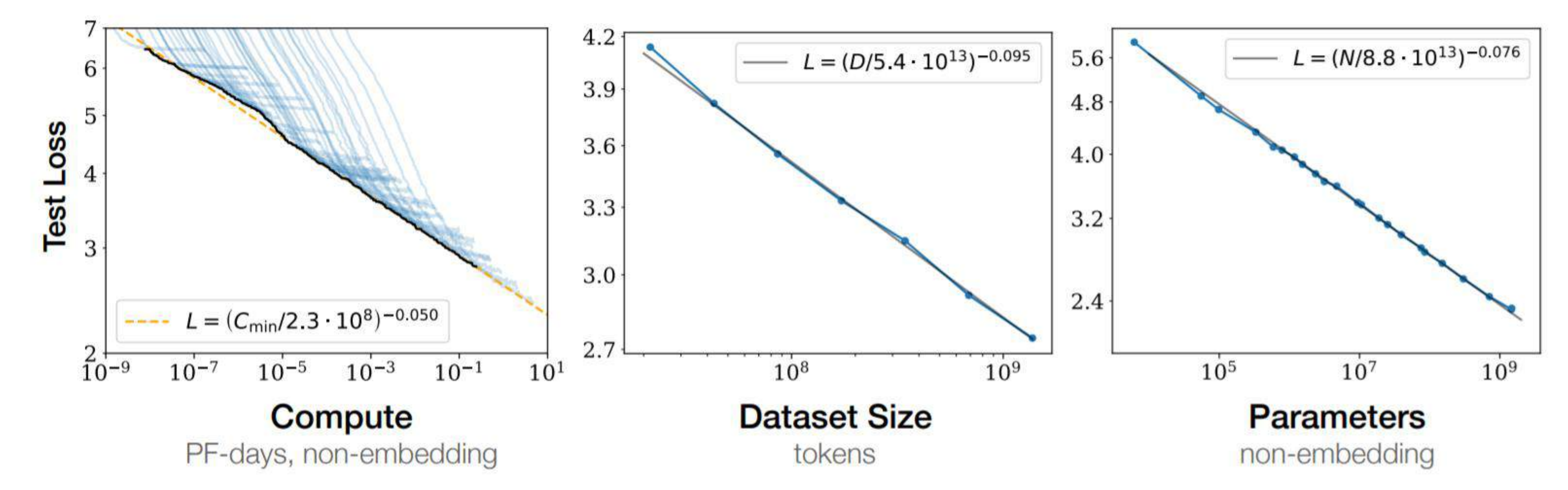
Input = *Where are we*
Output = *going*

Use-cases: auto-complete, spelling correction

Language Model

Baseline = text characteristics + data in abundance

Scaling Laws



https://arxiv.org/pdf/2001.08361.pdf (Scaling Laws for Neural Language Models)

” *Performance depends strongly on scale, weakly on model shape*”

Scaling Laws

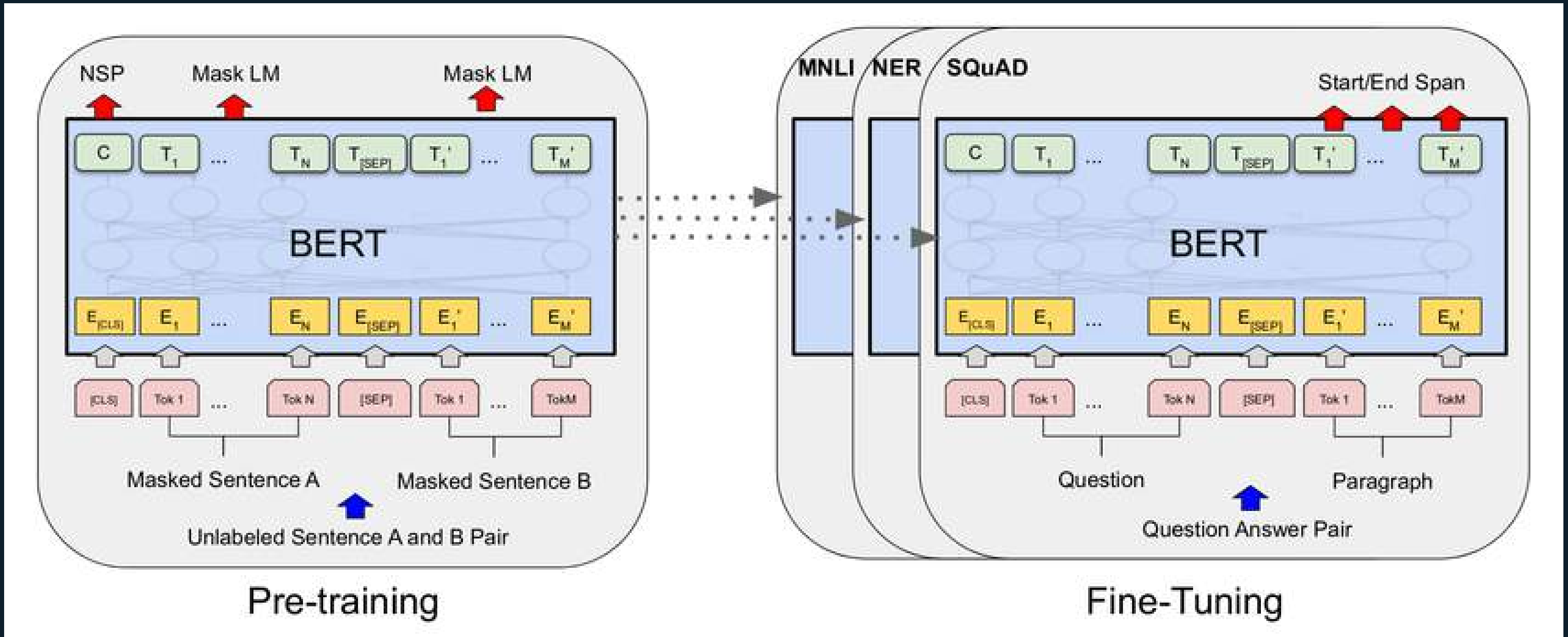
GPT3:

175B parameters

cost of 4.6 M€ as lower bound per training run; between 11.5 and 27.6M€ total development cost

AI SWEDEN

Adapting



First: Pre-Training LM

Then: Fine-Tuning (Transfer Learning) and/or prompting

Prompting

Prompting for Knowledge Acquisition

What does a GPT model know? [EleutherAI - text generation testing UI]

One-shot: New Delhi is the capital of

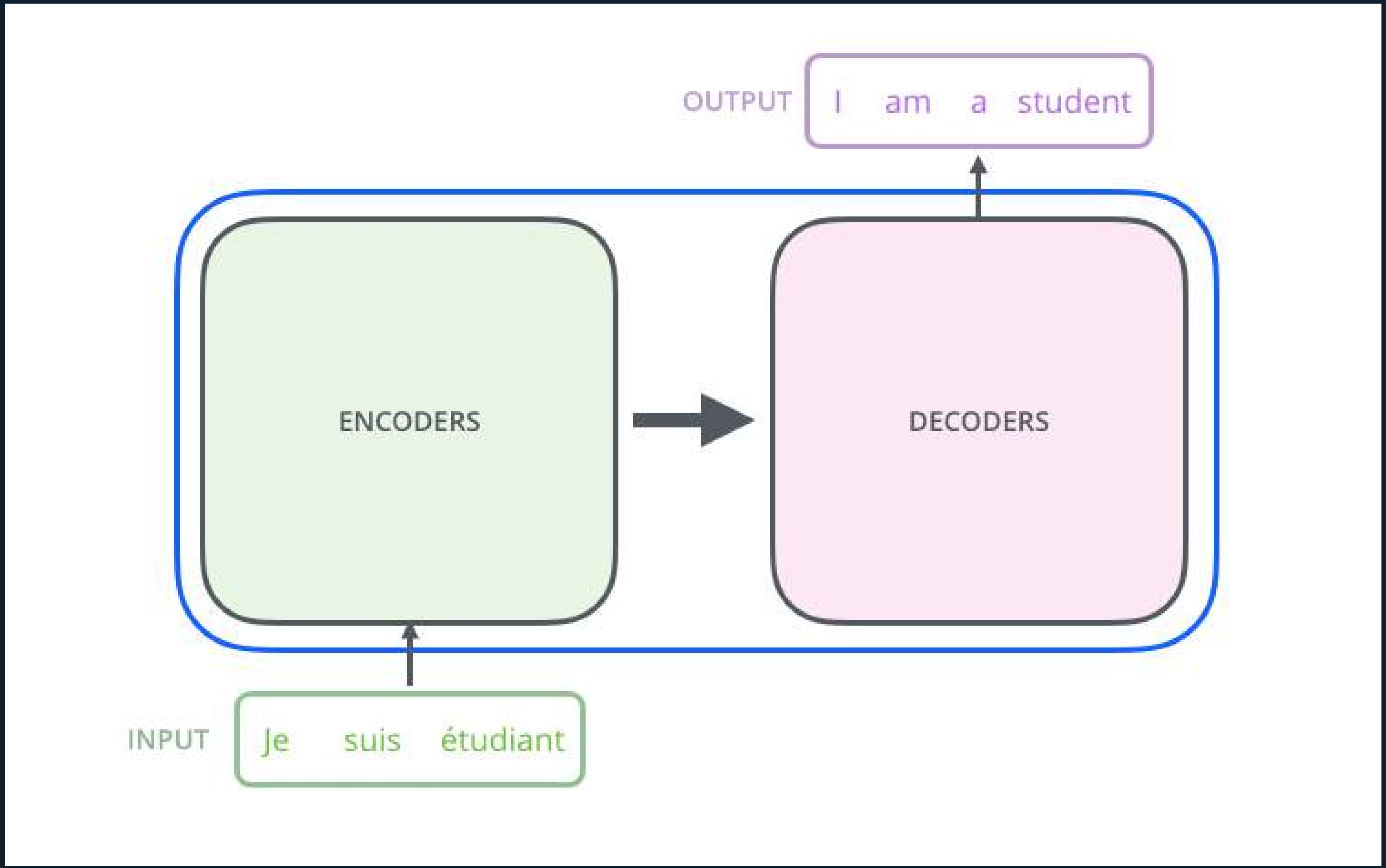
Few-shot:

1. Stockholm is the capital of Sweden.
2. Brussels is the capital of Belgium.
3. Taipei is the capital of Taiwan.
4. Kampala is the capital of Uganda
5. New Delhi is the capital of

Architecture

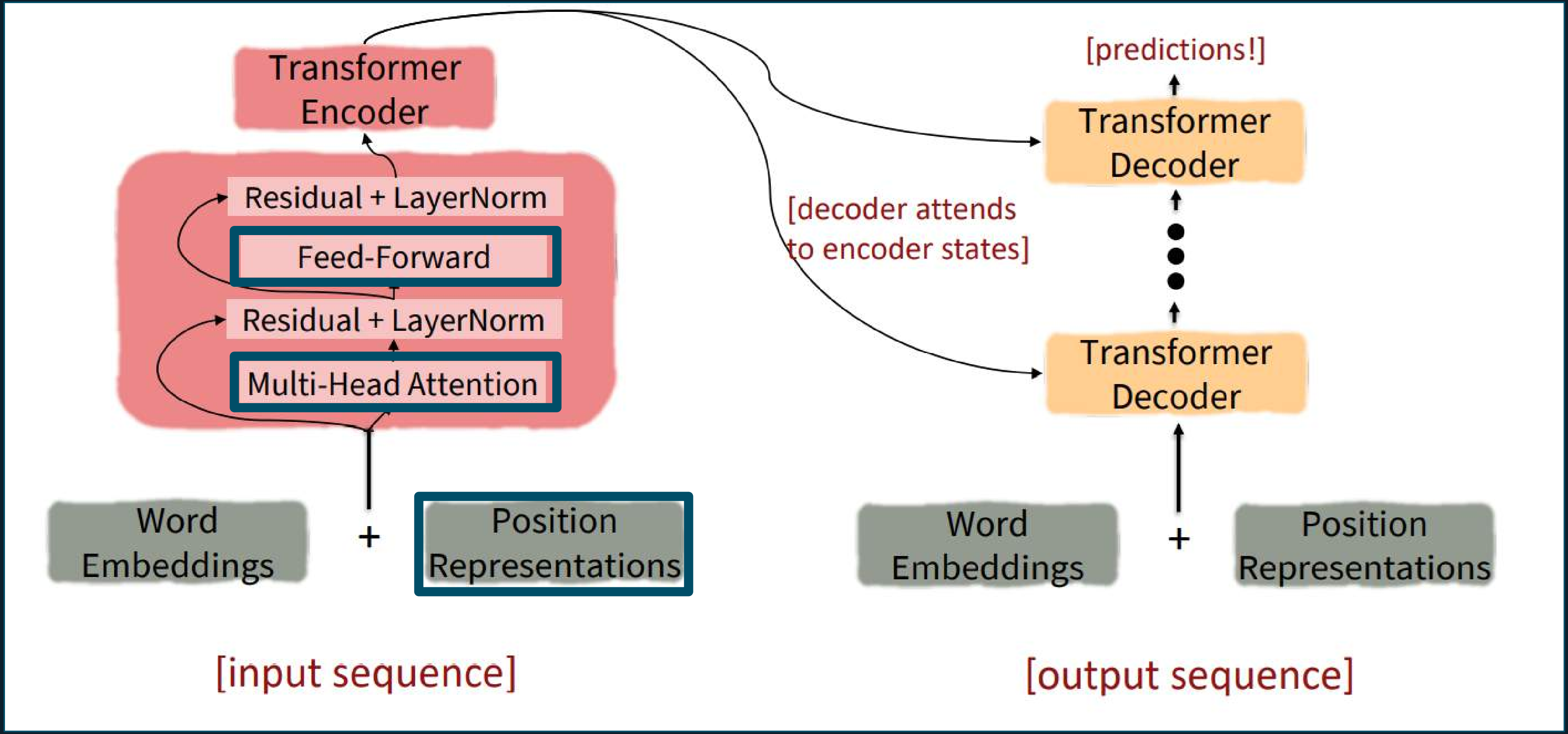
Attention

TRANSFORMERS

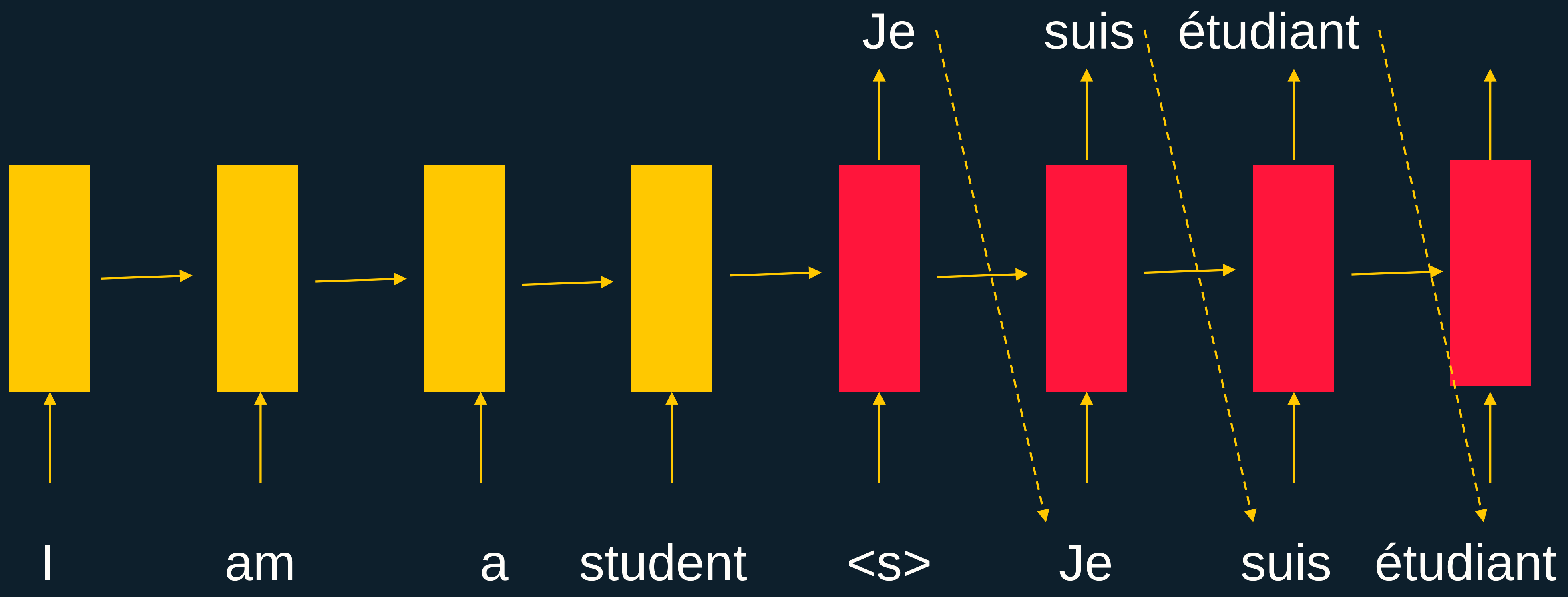


[The Illustrated Transformer – Jay Alammar – Visualizing machine learning one concept at a time. (jalammar.github.io)](jalammar.github.io)

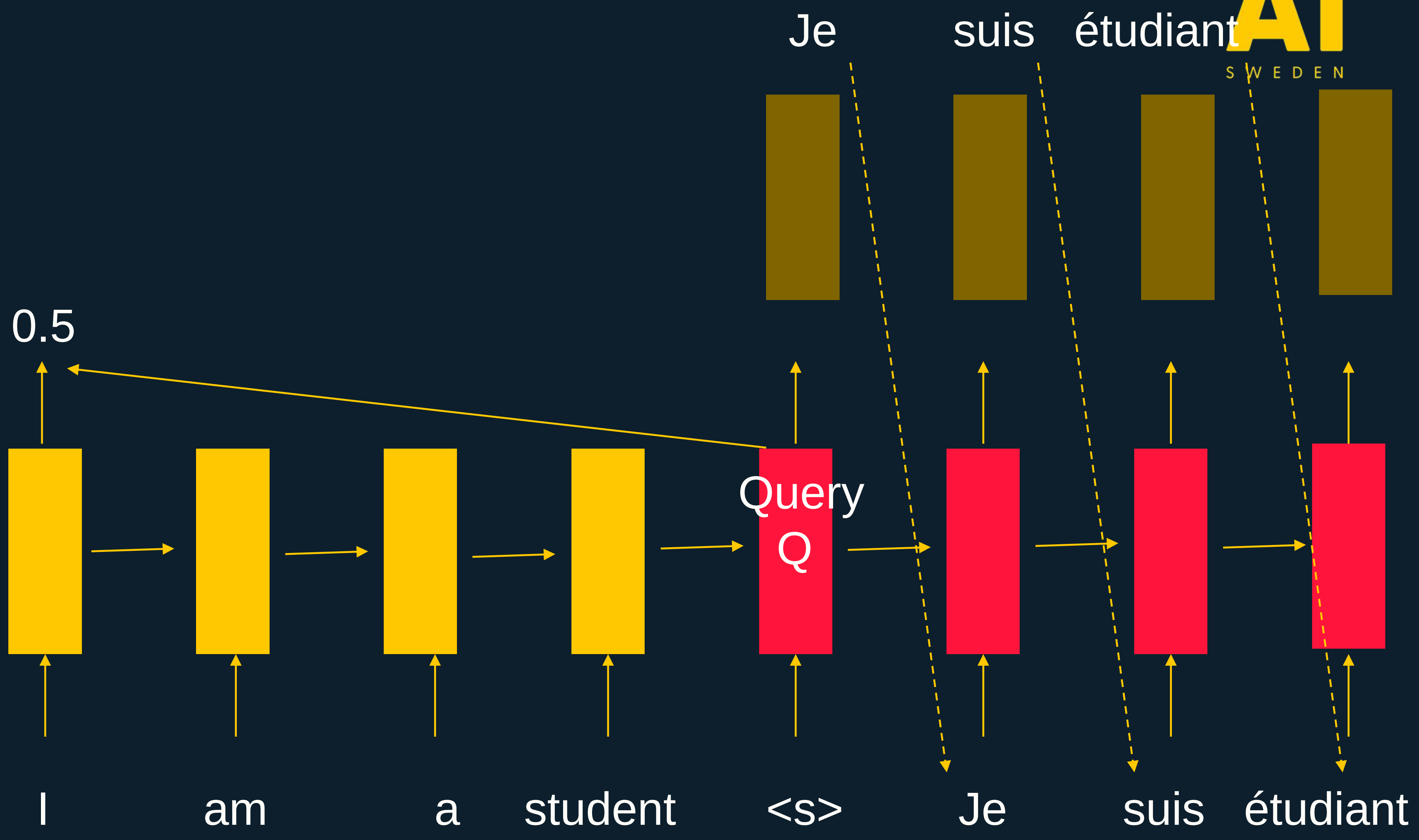
TRANSFORMERS



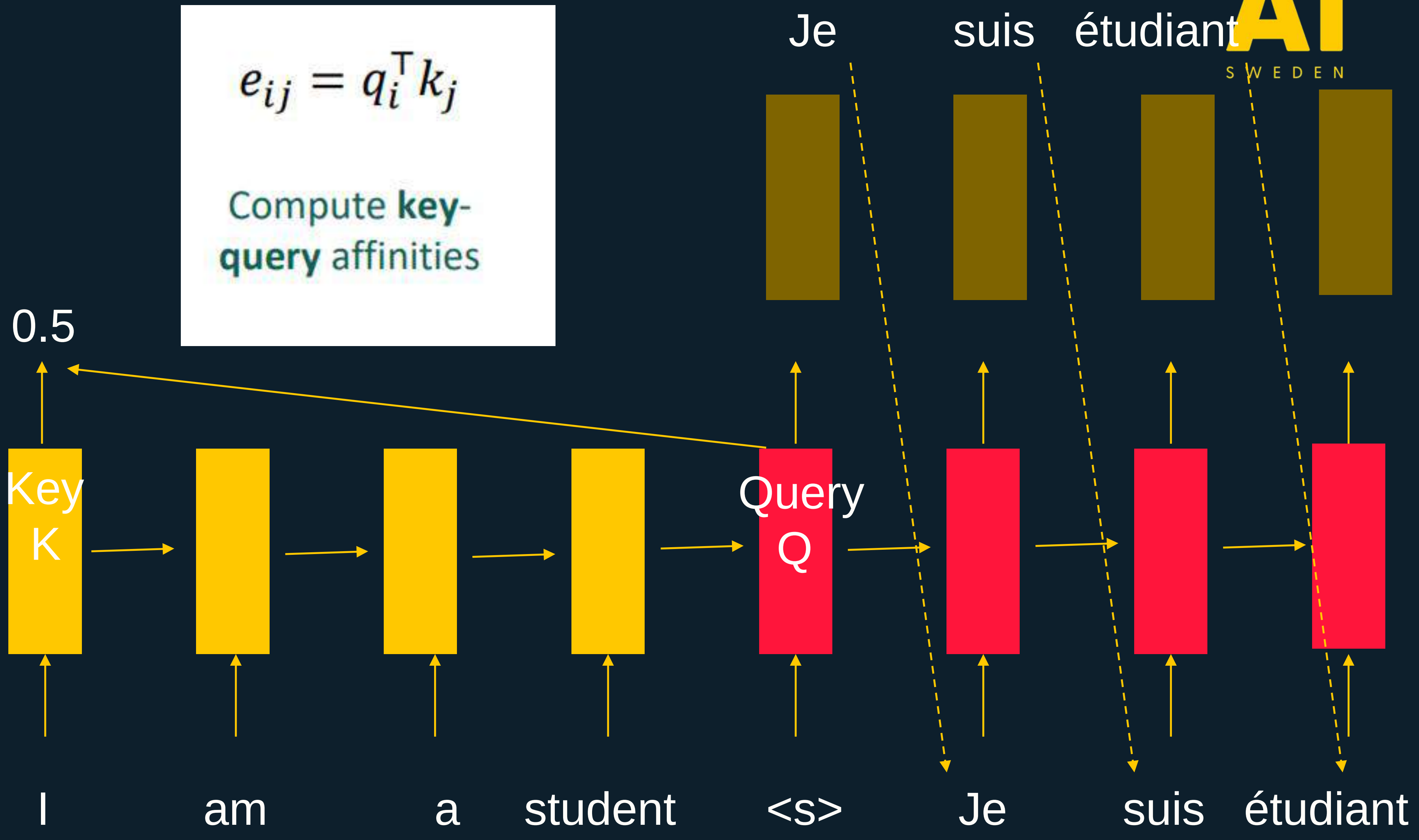
MACHINE TRANSLATION



ATTENTION



ATTENTION

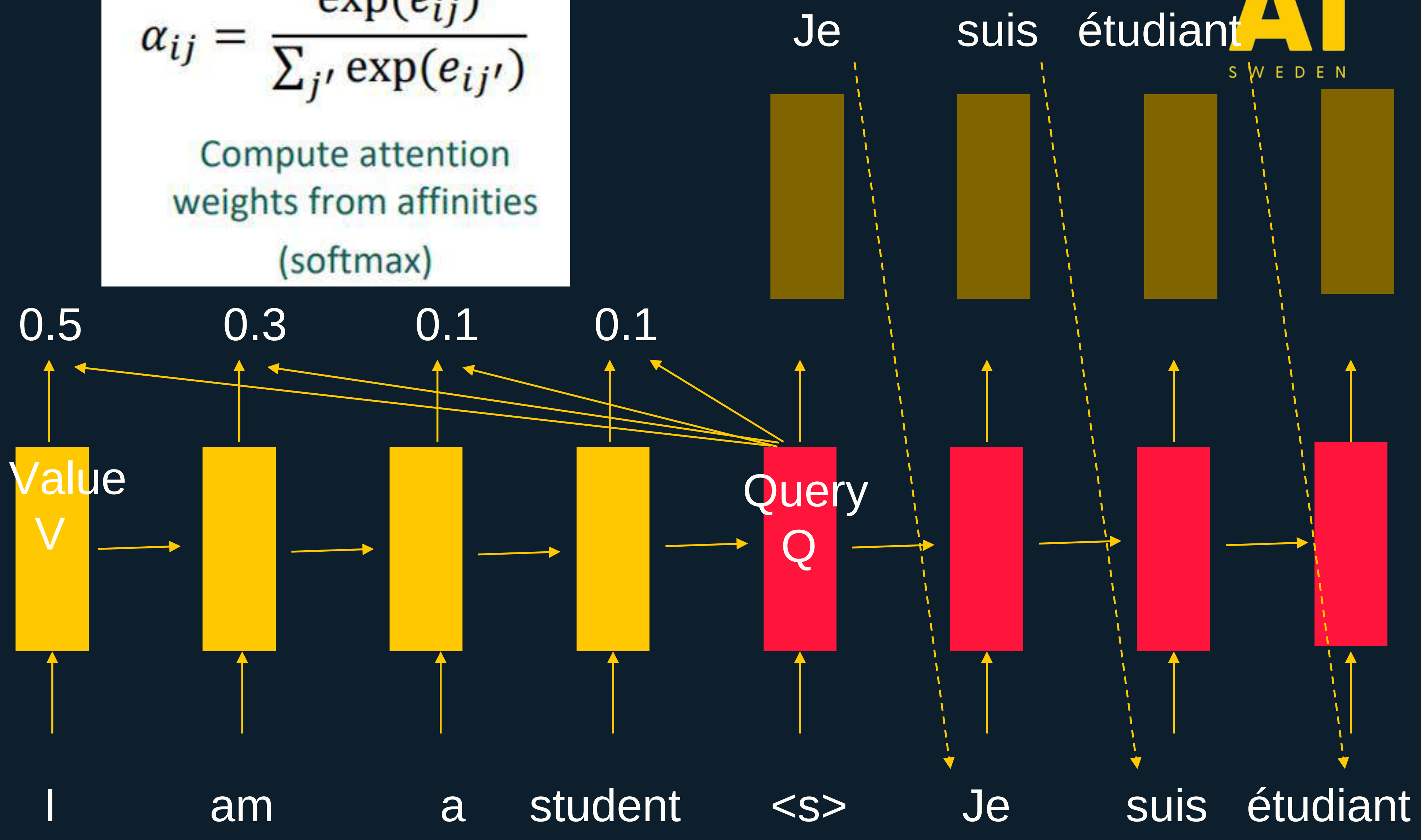


$e_{ij} = q_i^\top k_j$
Compute **key-query** affinities
0.5
Je
suis
étudiant
Key
K
Query
Q
I
am
a
student
<s>
Je
suis
étudiant

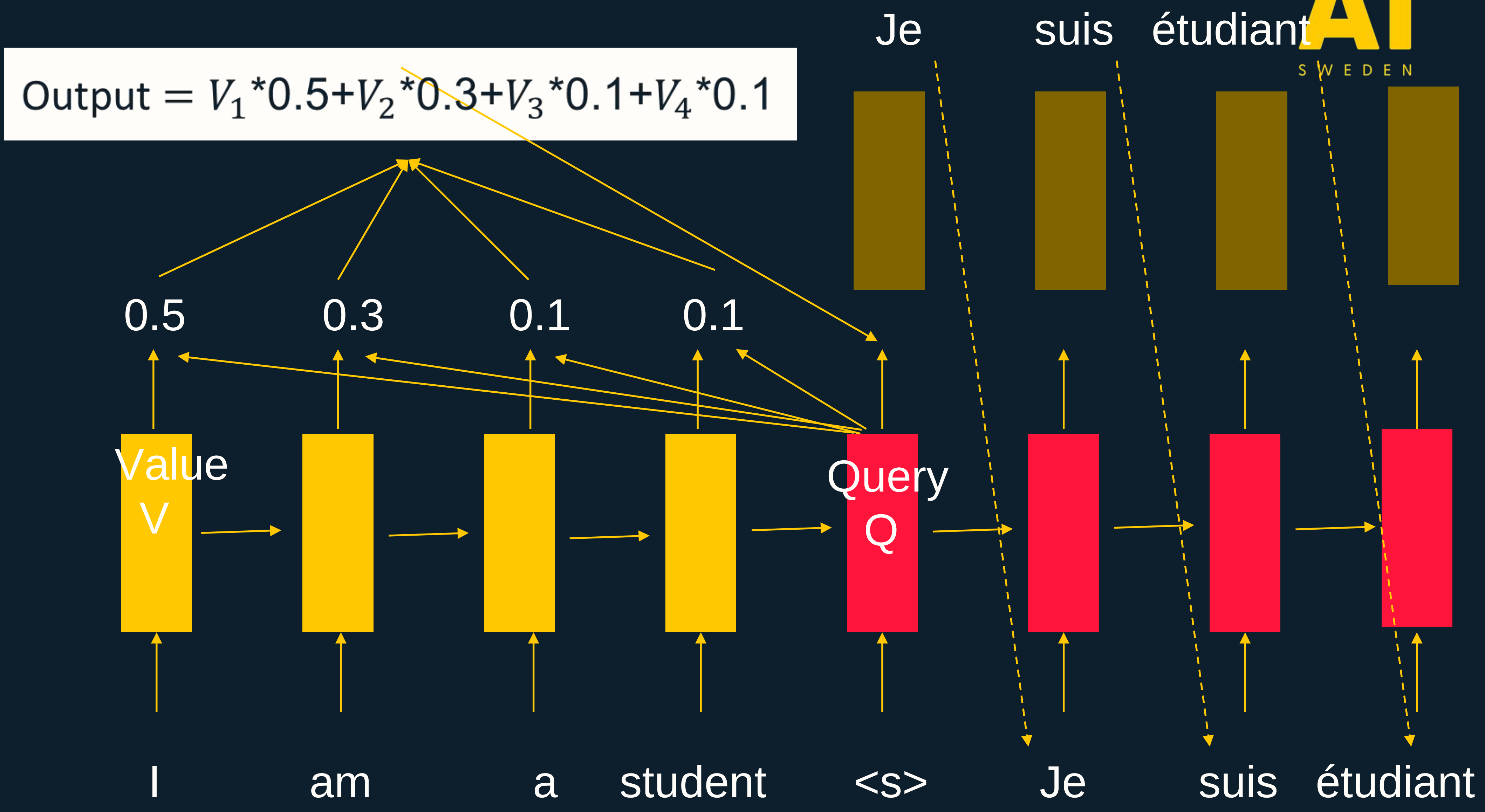
ATTENTION

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{j'} \exp(e_{ij'})}$$

Compute attention weights from affinities (softmax)



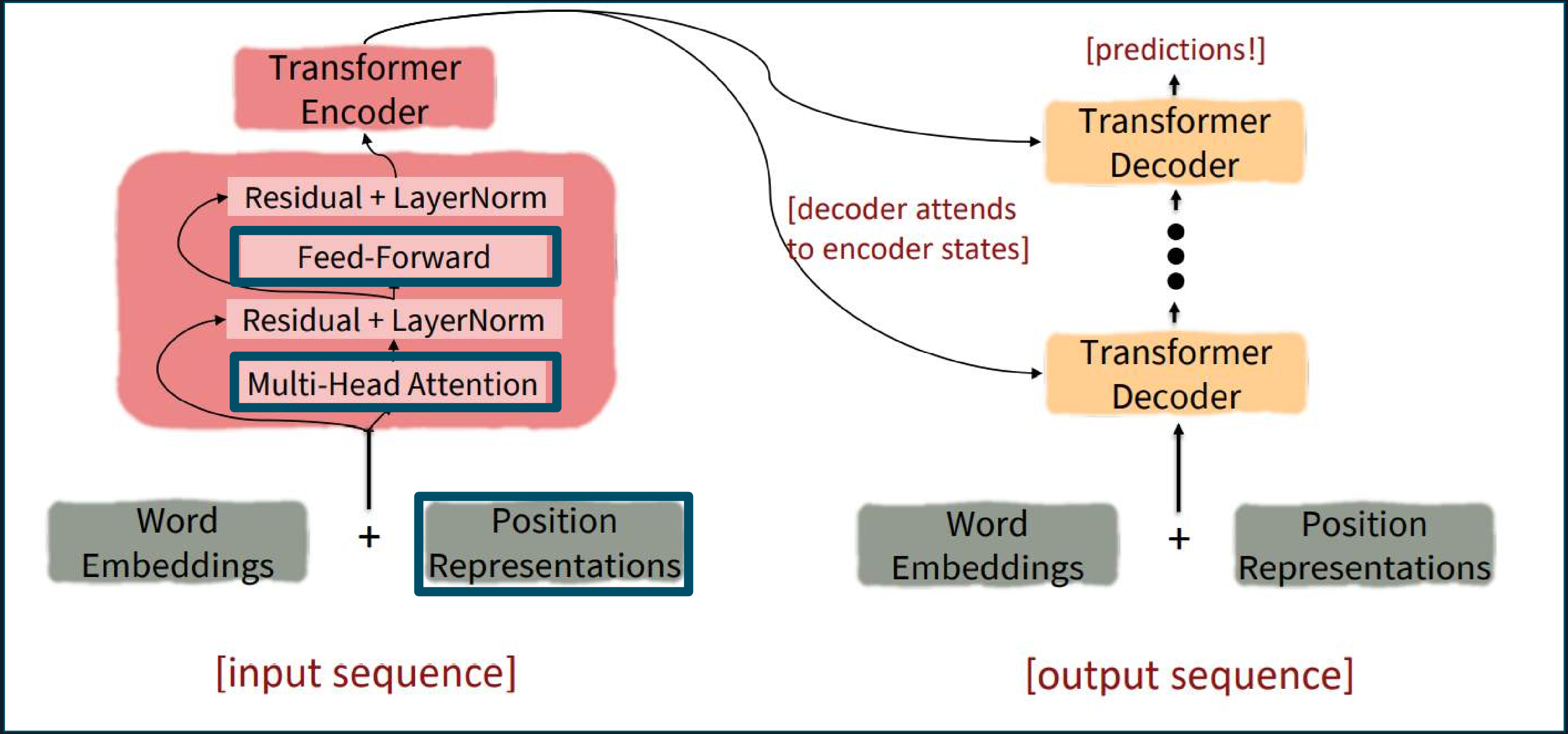
ATTENTION


Output = V_1*0.5+V_2*0.3+V_3*0.1+V_4*0.1
Je
suis
étudiant
0.5
0.3
0.1
0.1
Value
V
Query
Q
I
am
a
student
<s>
Je
suis
étudiant

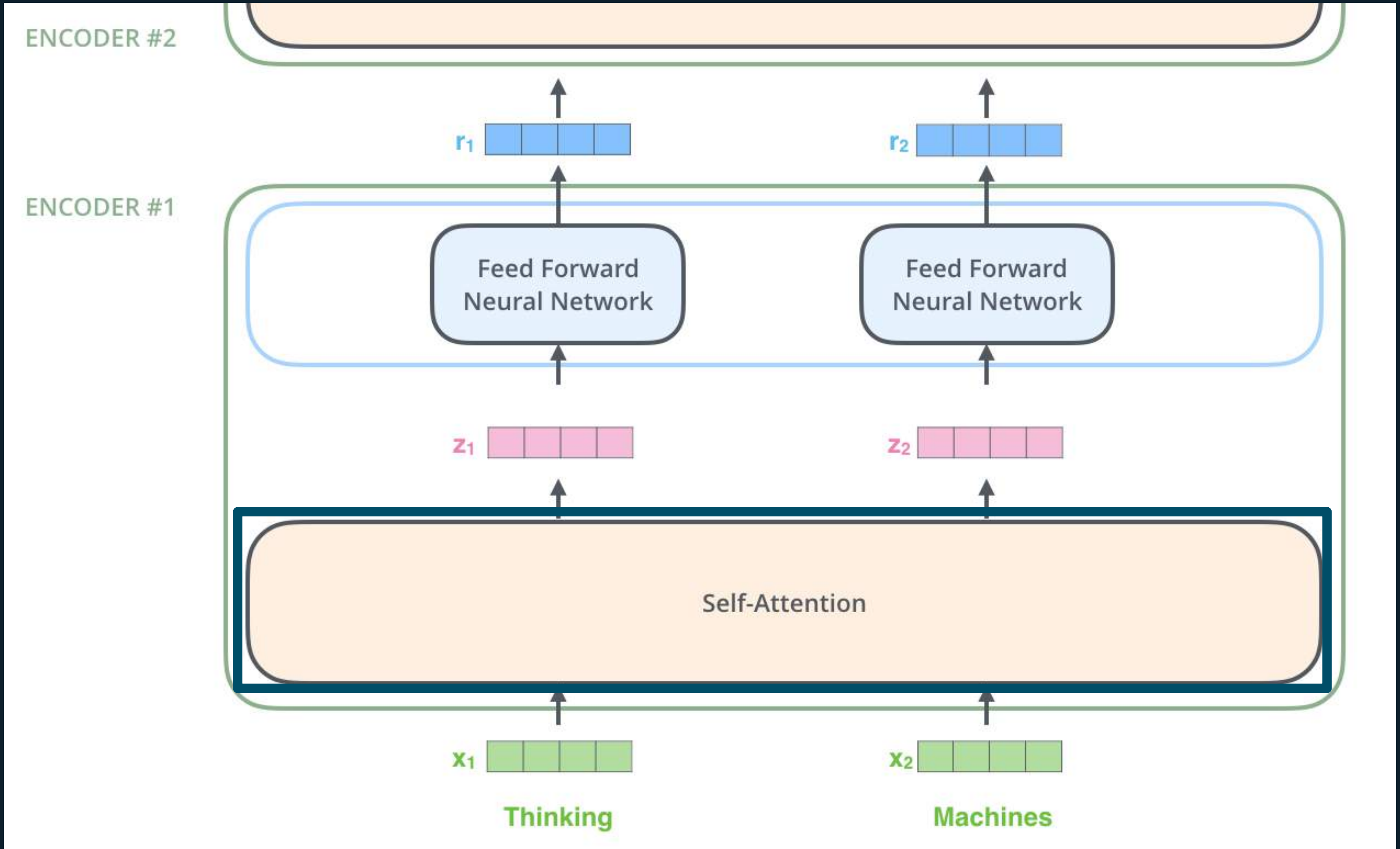
ATTENTION

= Long-Range Dependencies

TRANSFORMERS



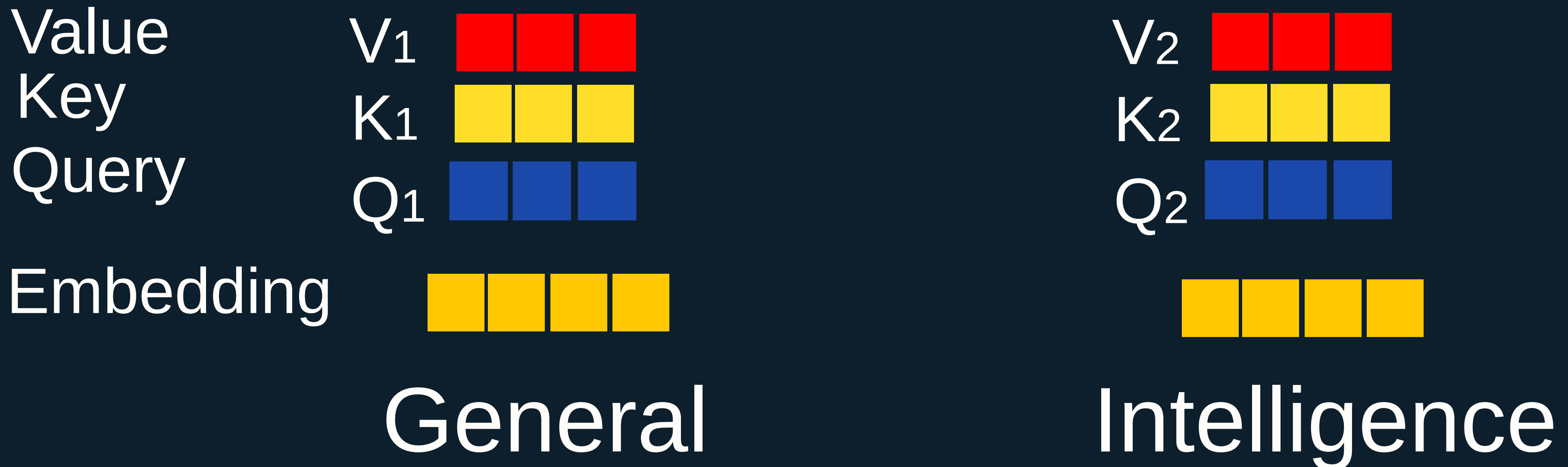
TRANSFORMERS



[The Illustrated Transformer – Jay Alammar – Visualizing machine learning one concept at a time. (jalammar.github.io)]

SELF-ATTENTION

SELF-ATTENTION



SELF-ATTENTION

Score

$Q_1 * K_1 = 112$

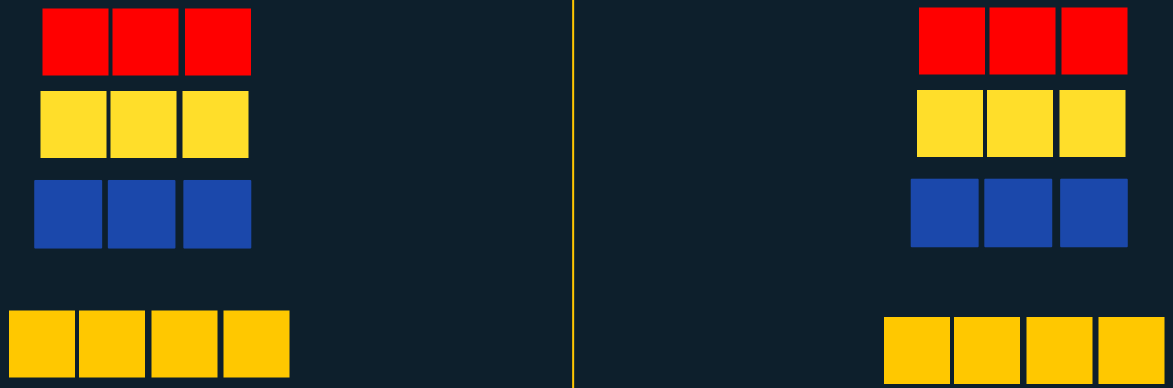
$Q_1 * K_2 = 96$

Value

Key

Query

Embedding



General

Intelligence

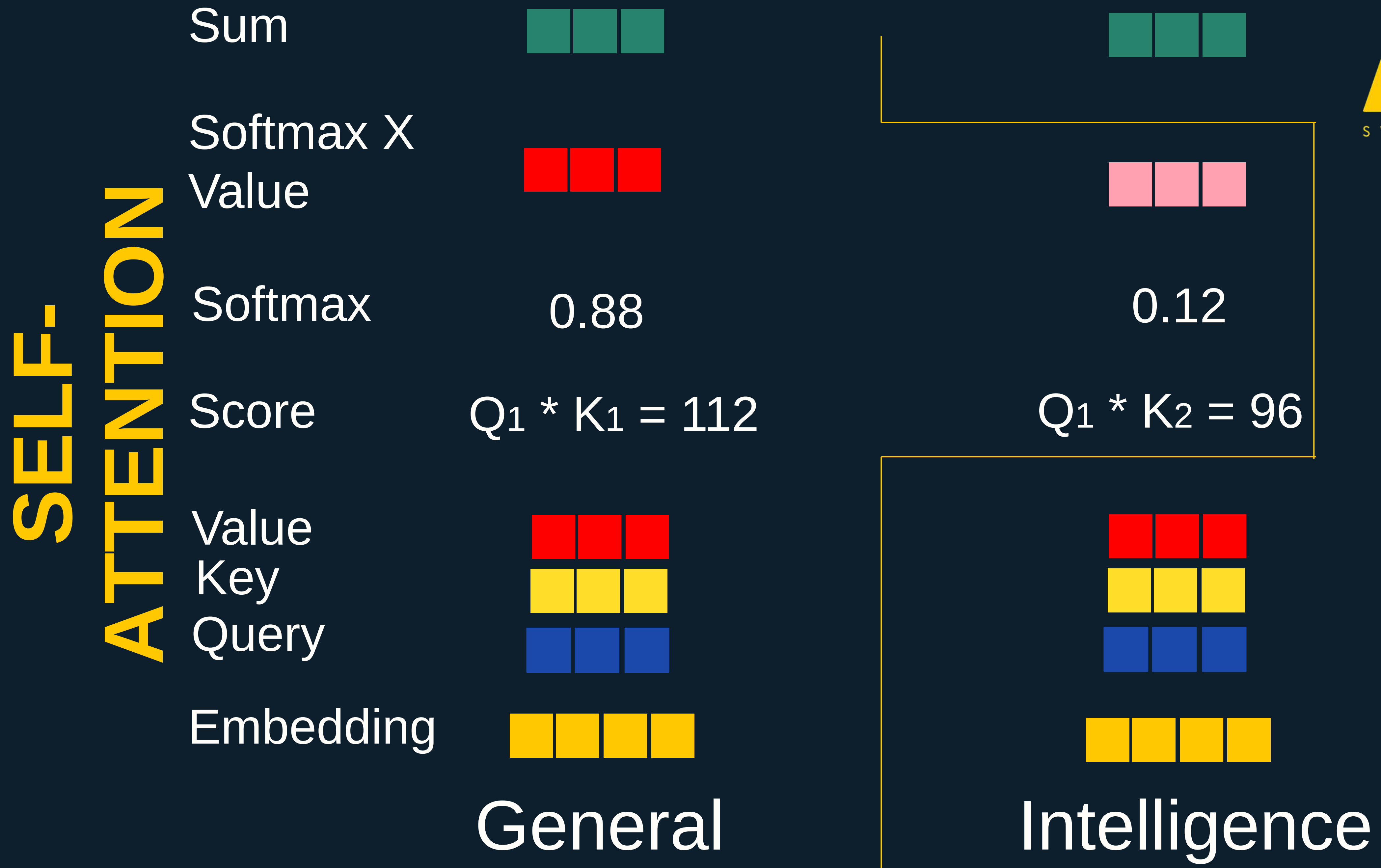
SELF-ATTENTION

	General	Intelligence
Softmax	0.88	0.12
Score	$Q_1 * K_1 = 112$	$Q_1 * K_2 = 96$
Value		
Key		
Query		
Embedding		

SELF-ATTENTION

AI SWEDEN

	General	Intelligence
Softmax X Value		
Softmax	0.88	0.12
Score	$Q_1 * K_1 = 112$	$Q_1 * K_2 = 96$
Value		
Key		
Query		
Embedding		

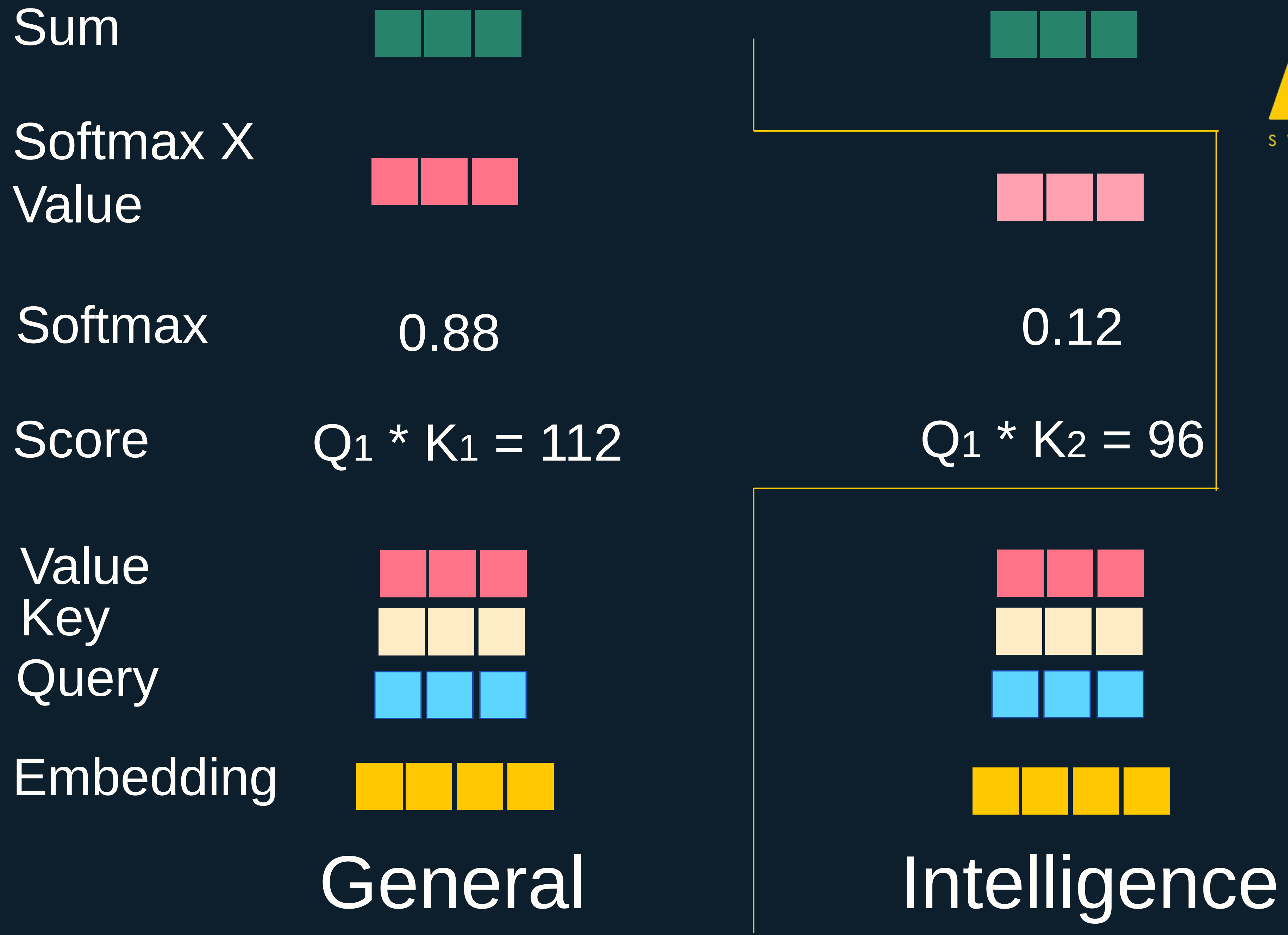


SELF-ATTENTION
Sum
Softmax X Value
Softmax
0.88
0.12
Score
Q1 * K1 = 112
Q1 * K2 = 96
Value
Key
Query
Embedding
General
Intelligence
AI SWEDEN

ATTENTION

= Omni-Directional

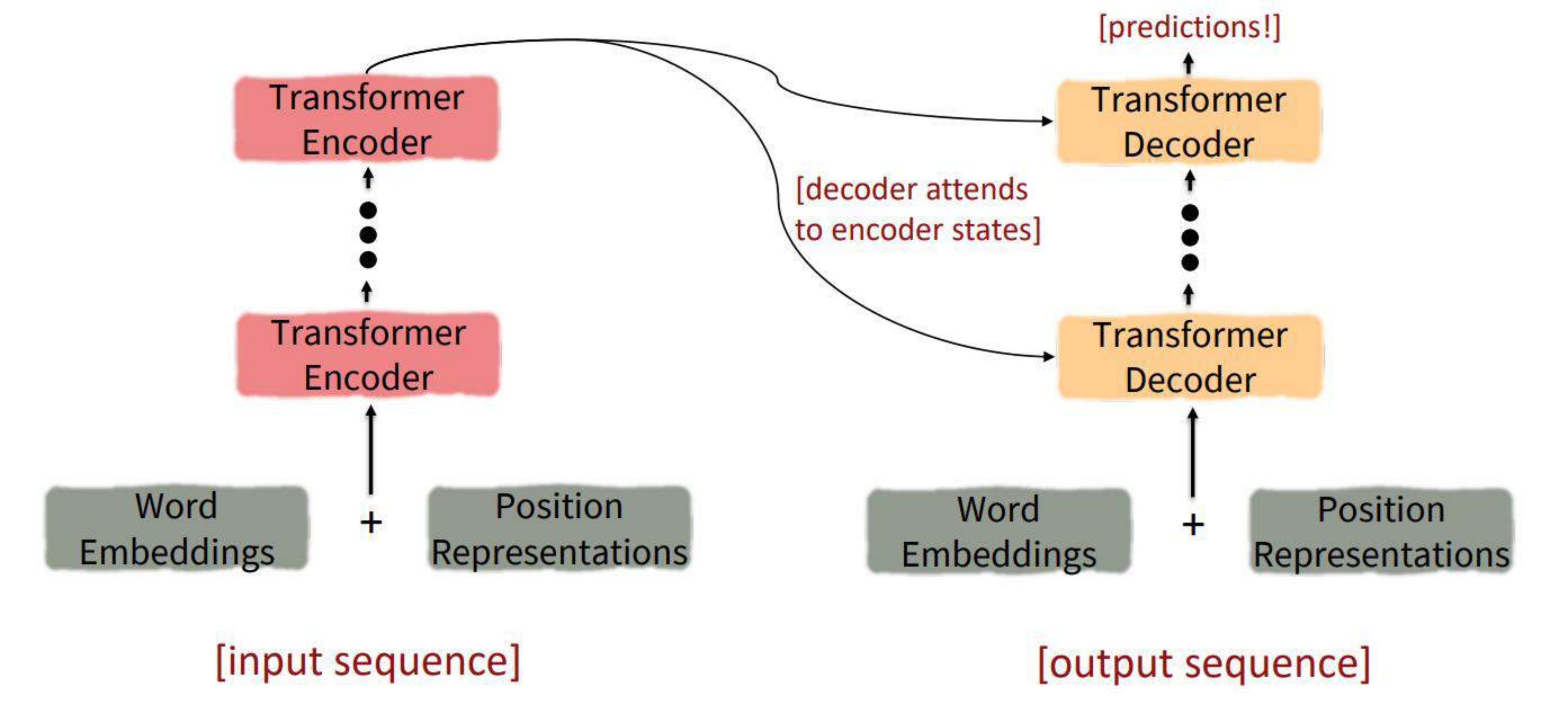
MULTI-HEAD



AI SWEDEN

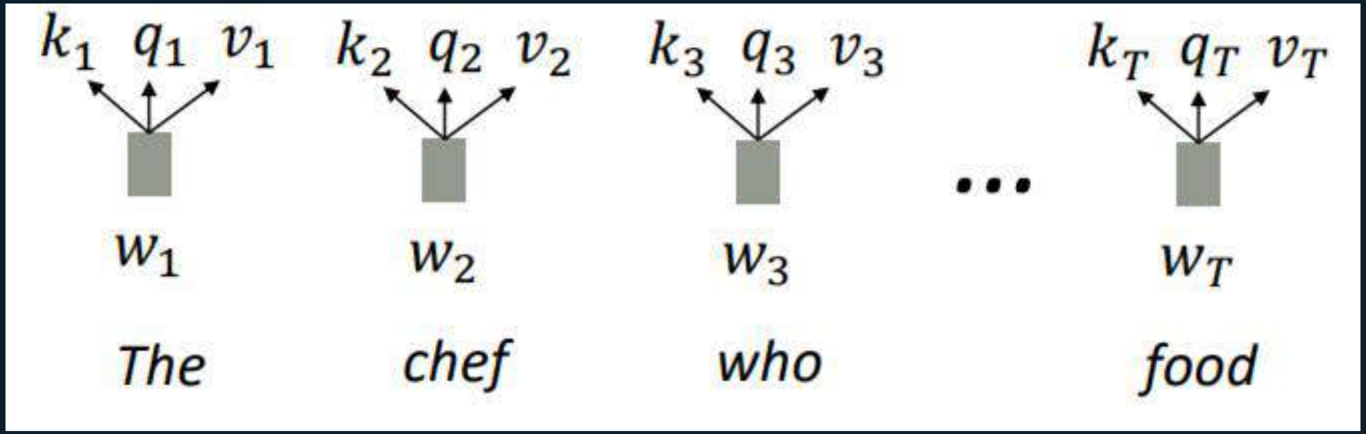
| Je ...

MASKS



I am a student |

POSITION EMBDEDDINGS

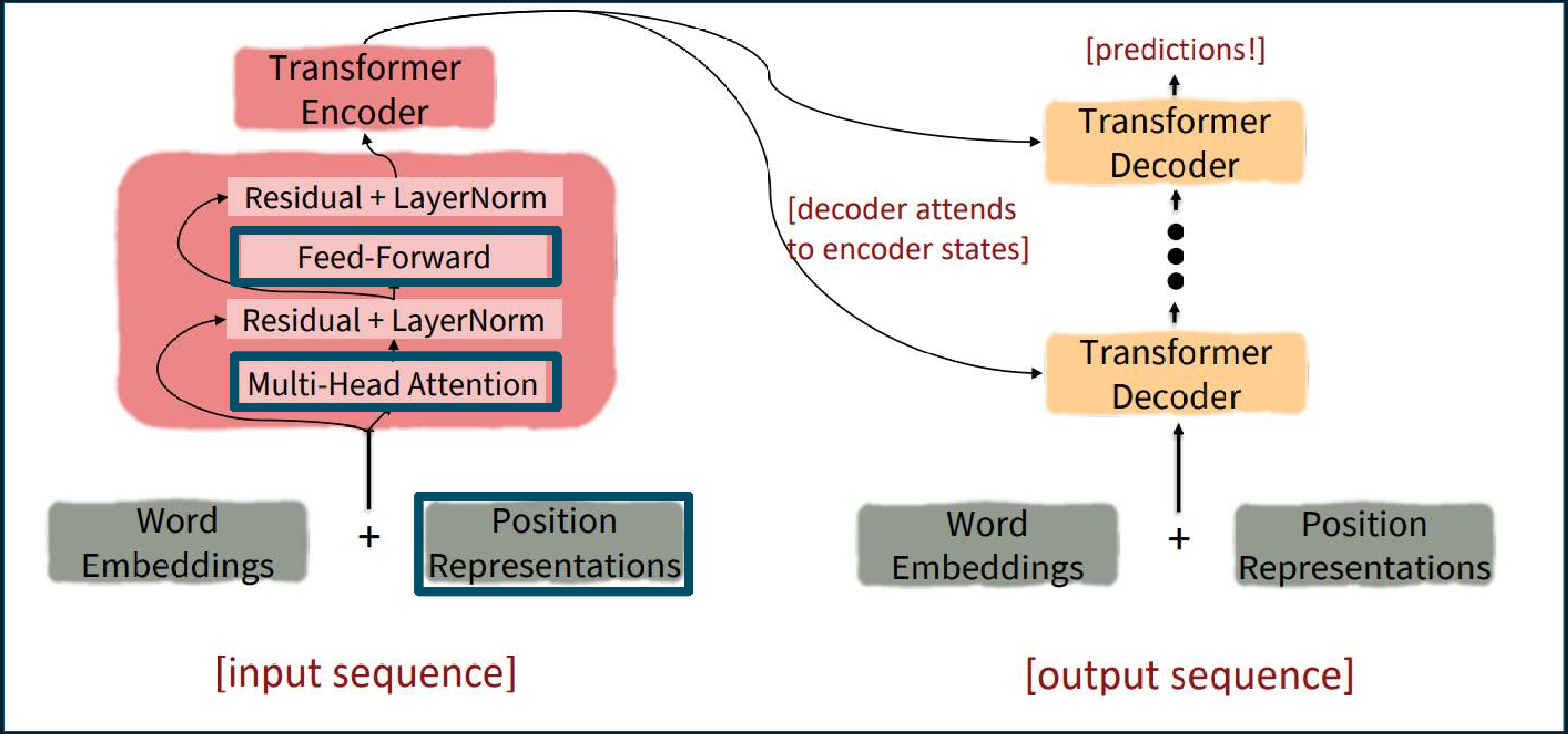


Order?

Postion Embeddings

- Learned embeddings
- Sinusoidal position representations

TRANSFORMERS



Types:

Sequence-to-sequence models

T5

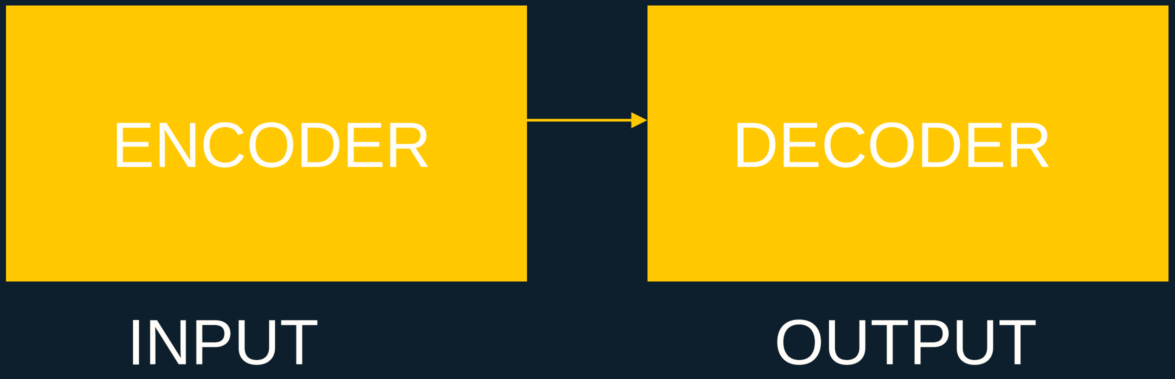
Encoders

BERT, roBERTa, KG-BERT, BART, …

Decoders

GPT-3, GPT-j, …

Seq-To-Seq



For example:

	input sequence	target sequence
machine translation:	*"translate English to German: That is good."*	*"Das ist gut."*
natural language inference:	*"mnli premise: I hate pigeons. hypothesis: My feelings towards pigeons are filled with animosity."*	*"entailment"* A single generated word per class label.
Winograd challenge:	*"The city councilmen refused the demonstrators a permit because *they* feared violence."*	*"The city councilmen"*

Goal: [CLS] my dog is cute [SEP] he likes play ##ing

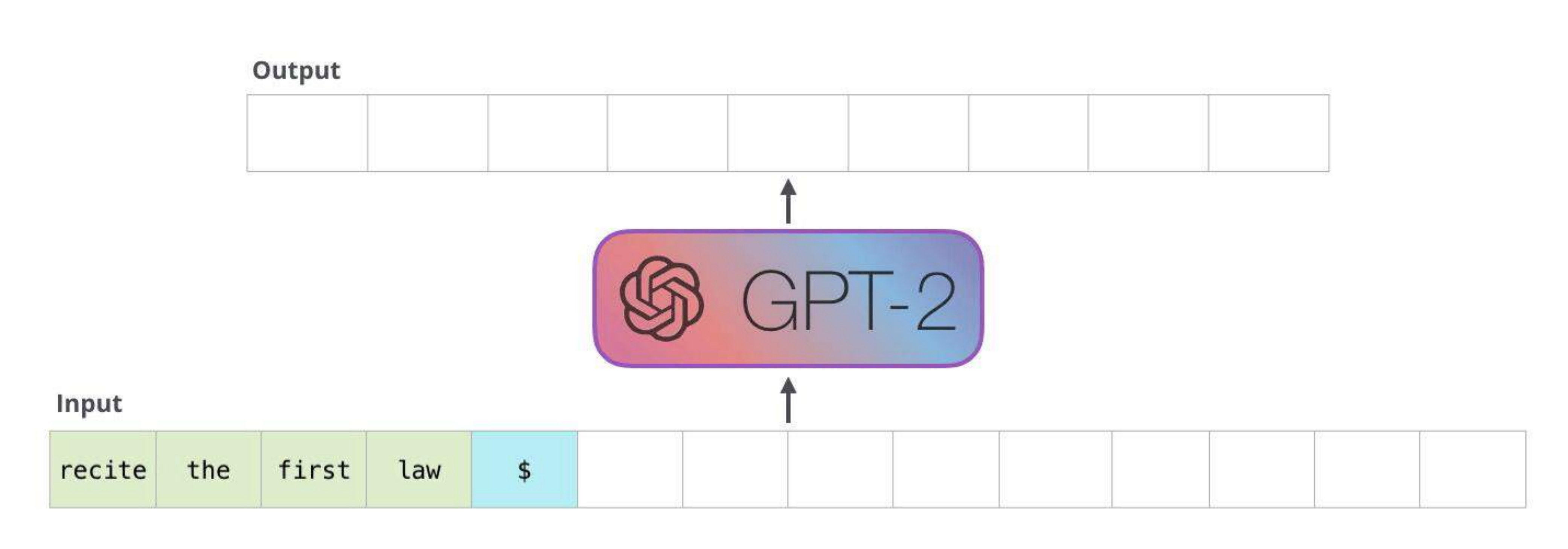
Encoder

1 (NSP) dog likes

Transformer

Input: [CLS] my [MASK] is cute [SEP] he [MASK] play ##ing

Decoder



[The Illustrated Transformer – Jay Alammar – Visualizing machine learning one concept at a time. (jalammar.github.io)]

AI
SWEDEN

SWEDISH NLP

Overview: Swedish NLP

Datasets & Evaluation

Text Corpora

Labeled datasets

SuperLIM

Models

Encoders

Decoders

Seq2Seq

Use Cases

Swedish Datasets - Corpora

- Swedish Wikipedia (~2GB)
- Litteraturbanken, Swedish literature (< 1GB)
- Oscar, crawled corpus filtered for Swedish (~30GB)
- Swedish Forums (> 10GB)
- National Library (KB), OCR & Radio etc. (~ a lot)
- Data Collection, work in progress...

Swedish Datasets - Labeled Data

- SUC 3.0, Swedish NER Corpus, Swe-NERC, Swedish Medical NER (PoS & NER)
- 🤗 Swedish Reviews (Sentiment Classification)
- Machine translated datasets (experimental)
- and more

Swedish Datasets - SuperLIM

- Swedish SuperGLUE evaluation suite
- Evaluates performance and bias

Swedish Datasets - SuperLIM

- Swedish SuperGLUE evaluation suite
- Evaluates performance and bias
- 13 test sets

Overview

Resource	Task
Aspect-Based Sentiment Analysis (Immigration)	Label the sentiment that the author of a text expressed towards immigration on the 1--5 scale
DaLAJ	Determine whether a sentence is correct Swedish or not
Swedish FAQ (mismatched)	Match the question with the answer within a category
SweSAT synonyms	Select the correct synonym or description of a word or expression
Swedish Analogy test set	Given two word pairs A:B and C:D, capture that the relation between A and B is the same as between C and D
Swedish Test Set for SemEval 2020 Task 1: Unsupervised Lexical Semantic Change Detection	Determine whether a given word has changed its meaning during a hundred year period
	Determine to what extent a given word has changed its meaning during a hundred year period
SweFraCas	Given the question and the premises, choose the suitable answer
SweWinograd	Resolve pronouns to their antecedents in items constructed to require reasoning (Winograd Schemata)
SweWinogender	Find the correct antecedent of a personal pronoun, avoiding the gender bias
SweDiagnostics	Determine the logical relation between the two sentences
SweParaphrase	Determine how similar two sentences are
SuperSim	Predict semantic word similarity and/or relatedness between words out of context.
SweWiC	Say if instances of a word in two contexts represent the same word sense.

https://spraakbanken.gu.se/en/resources/superlim

Swedish Models - Decoders

- GPT-SW3

Swedish Models - Seq2Seq

- Being explored

Swedish Models - Encoders

- AF-BERT (Swedish Public Employment Service)
- KB-BERT
- KB-ELECTRA
- KB-SBERT[1]
- Bigger & better encoders desired

[1]: https://kb-labb.github.io/posts/2021-08-23-a-swedish-sentence-transformer/

Swedish Models - Use Cases

Skatteverket: Anonymization

Severine likes good weather.

LINK: https://www.ai.se/sites/default/files/content/swebert_best_practices_2021-02-03.pdf

Swedish Models - Use Cases

Skatteverket: Anonymization

~~Severine~~ likes good weather.

LINK: https://www.ai.se/sites/default/files/content/swebert_best_practices_2021-02-03.pdf

FUTURE CHALLENGES

Future:

AI & ETHICS

Data based, Reflects bias in society

EVALUATION IN NLP

Does the results of the benchmark improve because of task understanding?

GENERALIZ ATION

GPT-3, GPT-j, …

Acknoledgements

Material based on

https://jalammar.github.io/illustrated-transformer/
https://jalammar.github.io/illustrated-bert/
https://jalammar.github.io/illustrated-gpt2/
https://jalammar.github.io/how-gpt3-works-visualizations-animations/

Material based on NLP course of Stanford Stanford CS 224N | Natural Language Processing with Deep Learning

Questions

1. How do you think that Attention/Transformers can be used in your working field?